STIC Search Report
Biotech-Chem Library

TO: Phuong Bui
Location: REM-2A15/2C18
Art Unit: 1638
Wednesday, March 10, 2004

Case Serial Number: 10/033109

From: Barb O'Bryen
Location: Biotech-Chem Library
Remsen E01A69
Phone: 571-272-2518

barbara.obryen@uspto.gov

Search Notes

RUSH

OM nucleic - nucleic search, using sw model

Run on: March 9, 2004, 10:35:31 ; Search time 6735 Seconds

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

(without alignments)
10978.948 Million cell updates/sec

Title: US-10-033-109-3
Perfect score: 1706
Sequence: 1 gcacgagcactcccaacccc........aaaaaaaaaaaaaaaaaa 1706
Scoring table: IDENTITY_NUC
               Gapop 10.0 , Gapext 1.0
Searched: 3470272 seqs, 2167151695 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   GenEmbl:*

```
 1: gb_ba:*
 2: gb_htg:*
 3: gb_in:*
 4: gb_om:*
 5: gb_ov:*
 6: gb_pat:*
 7: gb_ph:*
 8: gb_pl:*
 9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*
```

score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

c	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	910.8	53.4	1797	8	AF182188	AF182188 Lotus jap
	2	832.8	48.1	1557	8	AY135020	AY135020 Lotus jap
	3	768.6	45.1	1991	8	BT009318	BT009318 Triticum
	4	755.8	44.3	1506	6	AX412659	AX412659 Sequence
	5	755.8	44.3	1506	6	AX506508	AX506508 Sequence
	6	755.8	44.3	1700	8	ATAMT1	X75879 A.thaliana
	7	755.8	44.3	1748	6	A44314	A44314 Sequence 1
	8	755.8	44.3	1748	6	AR399484	AR399484 Sequence
	9	755.8	44.3	113970	8	AT6G15	AL049956 Arabidops
c	10	755.8	44.3	199634	8	ATCHRIV36	AL161536 Arabidops
c	11	753.8	44.2	69812	8	AP004059	AP004059 Oryza sat
	12	752.2	44.1	1853	8	AK107204	AK107204 Oryza sat
c	13	739.6	43.4	128913	2	AP004053	AP004053 Oryza sat
	14	738.2	43.3	2044	8	AK073718	AK073718 Oryza sat
	15	738.2	43.3	4123	8	AF289477	AF289477 Oryza sat
	16	738.2	43.3	103039	8	OSJN00068	AL606607 Oryza sat
	17	736.2	43.2	1751	8	LEAMT1	X92854 L.esculentu
c	18	730.4	42.8	1817	8	AK107601	AK107601 Oryza sat
	19	723	42.4	2987	8	AF289479	AF289479 Oryza sat
	20	720.2	42.2	1706	8	AY037219	AY037219 Arabidops
c	21	713.8	41.8	224188	2	AC139709	AC139709 Medicago
	22	705.8	41.4	1545	6	AX506795	AX506795 Sequence
	23	705.8	41.4	1642	8	AY093374	AY093374 Arabidops
	24	705.8	41.4	1846	8	AF110771	AF110771 Arabidops
	25	705.8	41.4	1847	8	AY062571	AY062571 Arabidops
c	26	705.8	41.4	118335	8	AC006193	AC006193 Arabidops
	27	700.4	41.1	1497	6	AX507192	AX507192 Sequence
	28	700.4	41.1	1671	8	AF083035	AF083035 Arabidops
c	29	700.4	41.1	80393	8	AP000382	AP000382 Arabidops
	30	699.4	41.0	1790	8	AF083036	AF083036 Arabidops
	31	694	40.7	1731	8	AF306518	AF306518 Brassica
	32	672.6	39.4	4654	8	AF289478	AF289478 Oryza sat
c	33	667.4	39.1	97155	8	ATF16A16	AL035353 Arabidops
	34	667.4	39.1	197655	8	ATCHRIV69	AL161573 Arabidops
	35	660.6	38.7	1768	8	AK105667	AK105667 Oryza sat
	36	660.2	38.7	127431	2	AC146819	AC146819 Medicago
c	37	635.6	37.3	139089	2	AC146571	AC146571 Medicago
	38	635.6	37.1	1858	8	LEAMT2GEN	X95098 L.esculentu
	39	632.6	37.1	2046	8	AF001505	AF001505 Oryza sat
	40	539.6	31.6	2720	8	LJA279059	AJ279059 Lotus jap
	41	527.0	30.9	1004	8	AF080541	AF080541 Nepenthes
	42	452.2	26.5	1659	8	AF118858	AF118858 Lycopersi
	43	300.2	17.6	555	8	AF188744	AF188744 Brassica
c	44	298.8	17.5	591	11	AL772895	AL772895 Arabidops
	45	265.2	15.5	468	8	AY113167	AY113167 Arabidops

ALIGNMENTS

RESULT 1
AF182188
LOCUS   AF182188   1797 bp   mRNA   linear   PLN 24-OCT-2000
DEFINITION   Lotus japonicus putative ammonium transporter AMT1;1 mRNA, complete cds.
ACCESSION   AF182188
VERSION   AF182188.1   GI:10952509
KEYWORDS   .
SOURCE   Lotus corniculatus var. japonicus (Lotus japonicus)
ORGANISM   Lotus corniculatus var. japonicus
              Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
              rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Loteae;
              Lotus.
REFERENCE   1 (bases 1 to 1797)

Pred. No. is the number of results predicted by chance to have a

```
AUTHORS     Simon,U. and Udvardi,M.K.
TITLE       cDNA encoding a putative ammonium transporter from Lotus japonicus
JOURNAL     Unpublished
REFERENCE   2  (bases 1 to 1797)
AUTHORS     Simon,U. and Udvardi,M.K.
TITLE       Direct Submission
JOURNAL     Submitted (01-SEP-1999) Udvardi, Max-Planck-Institute for Molecular
            Plant Physiology, Am Muehlenberg 1, Golm D-14476, Germany
FEATURES    Location/Qualifiers
   source   1..1797
            /organism="Lotus corniculatus var. japonicus"
            /mol_type="mRNA"
            /variety="japonicus"
            /db_xref="taxon:34305"
            /tissue_type="nitrogen-fixing nodules"
   CDS      76..1581
            /codon_start=1
            /product="putative ammonium transporter AMT1;1"
            /protein_id="AAG24944.1"
            /db_xref="GI:11952510"
            /translation="MALPECSAANLAQLIGPNATDAAAVAGFICDQFTAVGQRFSDTA
            FAVDSTYLLFSAYIVFPSMQLGFAMLCAGSVRAKNTMNIMLTNVLDAAAGLFYYLFGF
            APAFGAPSNGFIGRHFFGLKDVPTVAFDYSYFLYQWAFAIAAAGITSGSIAERTQFVA
            YLIYSSFLTGFVYPVSVSHWFWSGDGWASATNTGNLLFGTGVIDFAGSGVHMVGGVAG
            LWGALIEGPRIGRFDHAGRAVSLRGHSASLVVLGTFMLMFWGYNPGSSFTKILSAYD
            SGNYYGQWSAIGRTAVTTLAGCTAALTLFGKRIISGHWNVTDVCNGLLGGFAAITA
            GCSVVEPWAAIVCGFVAAVVLIACNMLAEKVRYDDPLEAAQLHGGCGAWGIIFTALFA
            KEEYVNQYYPGKPRPYGLFMGGGKLLGAHVIQILVIIGWVSAIMGPLPFILNRMKL
            LRISTEDELAGMDLTRHGFAYAYEDDESHKPGIQLRKIEPNSSTPSAES"
ORIGIN
Query Match    53.4%;   Score 910.8;   DB 8;   Length 1797;
Best Local Similarity  74.2%;   Pred. No. 1.8e-151;
Matches 1181;   Conservative   0;   Mismatches 402;   Indels   9;   Gaps   2;
```

```
Qy      20 CCACCCGTAGTTCTTACCACCTTCAGTCACGGCGTAATACACTAACCAACCCACCATGTC   79
           || || ||||||||| ||||   |||    |||  |||  ||||
Db      21 CGGCCTCTTGACTCTTGACTCTCCCTTCTCTTTCCAAGTCCACTACCCCAAAAACGATGGC   80

Qy      80 GCTGCCTCGTTGTCCGCCGAACAACTGGCCCAACTTCTCGGCCCCAAACACCACAGACGC  139
           ||||||      |||  |||||| |||||| ||| ||||||||||
Db      81 GCTGCCGGAGTGCTCGGCCGGCGGCGAACCTAGCGGCAGCCTCATCCGGCCGACGACGC  140

Qy     140 CTCCGCGCCGCGCTCCCTTATCTGCGGCCATTTGCCGCCGCCGTGGACAGCAAGTTCGTCGA  199
           |||  ||||||         ||||||| || |||
Db     141 CGGCGCAGTCGCCCGGATTCATCTGCGACCAGTTCACCGCCGTGGGGTCAACGATTCTCCGA  200

Qy     200 CACGGCCTTCGCCGTCGACAACACCTACCTCCTCTTTTCCGCCTACCTCGTTTTTTTCTAT  259
           ||| |||||||||||   |||||||||||||| |||  ||||||||||
Db     201 CACAGCATTCGCGTGGACAGCACCTACCTCCTCTTCTCCGCTTACATCGTGTTCCCGAT  260

Qy     260 GCAGCTCGGCTTCGCCATGCTCTGGGCCGGCTCCGTCGGCGGGCCGTCGCGCCAAGAACACCATGAACAT  319
           ||||| |||||||||||||||     ||||||||| |||  ||||||
Db     261 GCAGCTCGGCTTCGCCATGCTCTGCGCGGGCAGCGTCCGCGCCAAGAACACCATGAACAT  320

Qy     320 CATGCTCACCAACGTCCTGGACGTCCTGGACGCTGCCGCGGGCCTCTTCTACTACCTCTTCGGCTT  379
           |||||||| |||||       |||   | |||||||||||||   ||
Db     321 CATGCTCACCAACGTCCTGGACGCGGCCGCGGGCCTGTTCTATCTCTTCGGTTT  380

Qy     380 CGCCTTCGCCTTTGGCTTTCATCGGCTAAACATTTCTTCGGCCTCAA  439
           ||||||||||||||    ||| ||| |||||| |||||||||
Db     381 CGCCTTTGCCTTCGGCGCACCCTCCAACGGCTTCATCGGCCGCCACTTCTTCGGCCTCAA  440

Qy     440 GGACATCCCTCATCCTCCTACGACTACAGCTACTCCTTACCCAATGGGCCTTCGGCAT  499
           ||  ||||   |||  ||||  ||| |||||||||  |||||||||||||||||
Db     441 AGACGTCCCCACCGTCGCTTTCGACTACAGCTACTTCCTCTACCAGTGGGCGTTTGCCAT  500

Qy     500 CGCCGCCGCCGGCATCACCAGCGGAAGCATCGCCGAACGCACACAGTTCGTGGCCTATCT  559
           |||||||||||||      ||||| |||||||     || |||||||||||||
Db     501 CGCCGCCGCCGGGATCACCAGCGGCTCAATCGCTGAACGCACACAGTTCGTTGCCTACCT  560

Qy     560 CATCTACTCCTCCTTCCTCACCGGCTTCGTCTATCCGGTGGTCTCCCACTGGTTCTGGTC  619
           ||||| || ||||||  ||||||||||| ||
Db     561 CATCTACTCGTCGTTCCTCACCGGGTTTGTCTACCCGGTGTCTCCACTGGTTCTGGTC  620
```

```
Db     620 CCCAGACGGCTGGGCCTCTGCCTTCAAGATCACCAACGGCCTATTTCCACCGGCGTAAT  679
           ||||| ||||||  |||  |||||||  ||||   ||||||||||| ||||||||||
Qy     621 CGGCGACGGTTGGGCCAGCGCCACCAACACCGGGAACTTGCTGTTTGGCACCGGGGTGAT  680

Qy     680 AGACTTCGCCGGTTCCGGCGTAGTCCACATGGTCGGCGGAATAGCGGCCTATGGGGAGC  739
           |||||  |||||| ||||||||||| ||||| |||||||||||| | || |||||||||
Db     681 TGATTTTGCCGGCTCGGGTGTGTGTGGTTCACATGGTGGTGCGGGGTCTTTGCGGGCGC  740

Qy     740 GCTGATCGAAGGCCCAAGAATGGACGTTCGATCATCGCAGGACGAGCTGTGGCCTTGCG  799
           || |||   |||| ||  |||||||||||     |||||  |||||||| ||||||||
Db     741 GCTTATTGAGGGACCCAGGATTGGCCGGTTTGACCATGCCGGTCGAGCCGTTTGCCTTCG  800

Qy     800 AGGCCACAGCCGCGTCCTTAGTAGTCCTTGGGAACCTTCTTGCTTTGGTTCGGTTGTACGG  859
           |   ||| || |||||||||  |||||||||||| ||||||| || ||||||||||||||
Db     801 CGGACAACAGTGCTTCGCTTGTTGTGCTTGGTACCTTCATGCTCATGTTCTGGGGTTACGG  860

Qy     860 ATTTAACCCCGTTCATTAACAAAATCCTACTTACGGTAACTCAGGAAATTACTA  919
           ||    | | |||  || || || ||||   |||   |||||||||| ||||||
Db     861 TTTTAACCCCGGTTCGTTCACCAAGATTCTCAGCGCCTA---TGATTCAGGCAACTATTA  917

Qy     920 CGGTCAATGAGAGCGCCGGTTGGCAGAACCGCGGTCACCACTACCCTAGCGGGGTCAACAGC  979
           ||| |   ||  || |||| ||||  |||| ||   |||||| ||||||  |||||||| |
Db     918 TGGTCAGTGAGGCGCGCAATAGGAAGGACCGCCGCTAACGACAACATTGGCTGCTGCACGGC  977

Qy     980 TGCCTTGACCACGCTATTCGGTAAACGGGTGATATCCGGTGTATCCCGGCGGATGACCGATGT  1039
           |||| |||||| |||   |||  ||  |||    |||||| ||  |    |||||| ||| ||
Db     978 GGGCGCTGACCACCCCTTTTCGGGAAAGAATTATCTCCCGGCCACTGGTCGGAACGTCACCGACGT  1037

Qy    1040 CTGCAACGGCGTGTTGTAGGCGGTTTCGGCGGATAACAGCCGGTTGCTCCGTGGTTGAGCC  1099
           ||||||||||       ||| ||||||||| |||||| ||||||||||| |||||||||||
Db    1038 ATGCAACGGCTTACTCGGCGGTTTCGGCGGATCATCAACCGCCGGTTGCTCCGTGGGAGCC  1097

Qy    1100 ATGGGCAGCCATCGTATGCGGTTTGTTGCTTCTAATAGTATTAATAGCTTGCAACAAATT  1159
           || ||||||||       ||||||||||| || || ||||||| |||  ||||||   ||
Db    1098 CTGGGCAGCCATCGTGTCGTGCGCCTTCGTCGCGCAGTTGCTTATCGGTGCAACATGCT  1157

Qy    1160 AGCAGAGAAGTTAAGTTCGACGATCCTCTGGAGGCGCGCAGTTGCACGGTGGTGTG  1219
           |||| ||||| ||||   |  |||||||| || ||||||   ||||| |||| |||
Db    1158 CGCAGAAAAGTGCGTTACGTACGACGACCCCTTAGAGGCAGCTCAGCTTCACGGCGTTGCGG  1217

Qy    1220 CACGTGGGGGGTGATATTCACGGCCGTTGTTCGCCAAAAAGGAGTATGTGAAGGAGGTTTA  1279
           ||| |||||      |||  |||| ||| || ||||||||||  |||||||| ||  ||||
Db    1218 CGCGTGGGGGATCATATTCACCGCGGTATTCGCCAAGGAGTACGTGAACCAGGTTTA  1277

Qy    1280 CGGGGT------GGGGAGGGGCCACGGGTTGCTCATGGGGGGTGGTGGGAAGTTGCTGGC  1333
           |||||       | || |||    |||| ||||||||  ||||| | ||| |||||| ||
Db    1278 CCCGGGTAAACCGGGTCGCGGCCATACGGGTTTCATGGGTGTGGGTGGAGGGAAGCTGCTCGG  1337

Qy    1334 GGGCGCACGTGATTCAGATTCGTGGTGATTGCTGTGGTTAGTGCGGCCATGGGAGACCCTT  1393
           |||||| | ||||| |||  ||||| | |||| ||||  ||||||||  |||||||||
Db    1338 AGCGCACGTGATCCAGATTTGGTGTGATATCGGTGGTCAGCGCCCAAATGGGTCCTCT  1397

Qy    1394 GTTTTGGGGGTGGTGAATAAACTGAAGCTGTTGAGGATGATTCTTCAAGGATGATGAGCTT  1453
           |||  |||||       | || ||| ||||||  |||| ||||| ||||| |||||||
Db    1398 GTTCTTCATTCTGAATAAGATGAAGTTGCTTCGGATCTCAACGGAGGATGATGAACTTGCGGG  1457

Qy    1454 GATGGACATGACTCGCCATCGAGAGGCTTATGCTTATGAGGATGATGAGACGCACAA  1513
           ||| || | | |||| |||| |  ||| ||| ||  || |||  |||| ||| ||
Db    1458 TATGGATCTGACCCGACCCGACATGGTGGGTTTGCATAGCATATGAGGATGATGAATCGCACAA  1517

Qy    1514 GCATGGGATGGCAGTTGAGGAGGGTTGGGCCCAACGCGTCTTCCACACCCACCACTGATGA  1573
           ||| | |||||||  ||||| |||   ||||||||| ||| |||| || |||  ||
Db    1518 GCCAGGGATTCAGCTACGGAAGAATTGAACCTAACCTAACTCTTCTTCCACTCCTAGTGCTGAGTC  1577

Qy    1574 ATGATCTTTTTTTCCCATATGCATGTCTCATT  1605
           ||||||| |      ||||| ||| ||| |||
Db    1578 ATGATAATAATTAGTAGTAATAAATTTCGTTTT  1609
```

```
RESULT 2
AY135020
```

```
LOCUS       AY135020              1557 bp    mRNA    linear   PLN 01-JUN-2003
DEFINITION  Lotus japonicus ammonium transporter (Amt1;2) mRNA, complete cds.
ACCESSION   AY135020
VERSION     AY135020.1  GI:31322043
KEYWORDS    .
SOURCE      Lotus corniculatus var. japonicus (Lotus japonicus)
  ORGANISM  Lotus corniculatus var. japonicus
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Loteae;
            Lotus.
REFERENCE   1  (bases 1 to 1557)
  AUTHORS   El Alaoui,H. and Chiurazzi,M.
  TITLE     Characterization of a new ammonium transporter in Lotus japonicus
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1557)
  AUTHORS   El Alaoui,H. and Chiurazzi,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JUL-2002) Plant Biology, Institute of Genetics and
            Biophysics 'Adriano Buzzati Traverso', Via Marconi,12, Naples
            80125, Italy
FEATURES             Location/Qualifiers
     source          1..1557
                     /organism="Lotus corniculatus var. japonicus"
                     /mol_type="mRNA"
                     /variety="japonicus"
                     /db_xref="taxon:34305"
                     /ecotype="GIFU"
     gene            1..1557
                     /gene="Amt1;2"
     CDS             1..1557
                     /gene="Amt1;2"
                     /note="putative AMT1;2"
                     /codon_start=1
                     /product="ammonium transporter"
                     /protein_id="AAM95453.1"
                     /db_xref="GI:31322044"
                     /translation="MASLSCSATDLAPLITATTNATATAAATYLCNQLDTISRKLSDT
                     TYAVDNTYLLFSAYLVFAMQLGFAMLCAGSVRAKNTMNIMLTNVLDAAAGGLSYYLFG
                     FAPAFGAPSNGFIGRHFFGLKHYPSPTYDYSFFLYQWAFAIAAAGITSGSIAERTQFV
                     AYLIYSSFLTGFVYPVVSHWLWSSDGWASPTRTGSLLFGSGAIDFAGSGVVHMVGGI
                     AGLWGAFIEGPRIGRFDRTGRSVALRGHSASLVVLGSFLLWFGWYGFNPGSFLTIAKA
                     YGNNGENGNYYGGWSAIGRTAVTTLLAGCTAALTTLFSKRLLEGHWKVLDVCNGLLGG
                     FAAITSGCSVVBPWPAIVCGFVAAWVLIGLNLVAAKMKYDDPLEAAQLHGGCGAWGVL
                     FTGLFAKGEYVAEVYGSAGRPYGLLMGGGKLLAAQVIEILVVCGWVTATMGPLFYGL
                     HKTKLLRISEDDETAGMDLTRHGFAYAYHDDDDVSTKRGVMMSRIGPGSSSPSTMNT
                     PAASAAANDC"
ORIGIN
```

```
Query Match      48.8%;  Score 832.8;  DB 8;  Length 1557;
Best Local Similarity   74.0%;   Pred. No. 1.2e-137;
Matches 1117;  Conservative  0;  Mismatches 372;  Indels 21;  Gaps 4;
```

Qy   378 TTCGCCTTCGCTTCGGC... TCCCCTCCAACGGCTTCATCGGTAAACATTCTTCGGCCTC 437
          ||||||||||||||||||...||||||||||||||||||||||||||||||||||||||||||
Db   307 TTCGGTTCGCCTTCGGCAGCCCCCTCCACGAGCCCCTCCACCGGCTTCATCGGCCGCCTC 366

Qy   438 AAGGACATCCCTCATCCTC...TACTCCTCCTCCACCAGGGCCTTCGCC 497
          ||||||||...|||||||||||||||||||||||
Db   367 AAACACTACCCTTCGCCCACCTACAGCTTCTTCCTCCTCCACCAGTGGCCTTCGCC 426

Qy   498 ATCGCCGGCCGGCCATCACCAGGGAAGCATCGCCGAACGCACACAGTTGTGGCCTAT 557
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   427 ATCGCCGCTGCCGGAATCACCAGCGGCTCCATCGCCGAGAGGACACAATTCGTCGCCTAC 486

Qy   558 CTCATCTACTCCTCCTTCACCGGCTTCGTCATCCGGTGGTCCCCCACTGGTTCTGG 617
          ||||||||||||||||...|||||||||||
Db   487 CTCATCTACTCCTCCTTCCTAACCGGCTTGTCTACCCGTCGTCGCACTGGCTCTGG 546

Qy   618 TCCCCAGACGGCTGGGCCTTCGCCTTAAGATCACCGACCG--GCTATTTTCCACCGGC 674
          |||||||||||||||... ||||| |||||||||||||
Db   547 TCCTCGGACGGCTGGGCCAGCCCACCCGGTCGACCACTGGCAGCTACTACTATTCGGATCCGGC 606

Qy   675 GTAATAGACTTCGGCCGGTTCCGGCGTAGTCCACATGGTCGGCGAATAGCCGGCTATGG 734
          | ||||||||||||||||||| |||||||||||||||||||||||
Db   607 GCGATTGGATTTCGCTGGATCCGGAGTAGTCCACATGGTCGGTGGCATAGCAGGACTATGG 666

Qy   735 GGAGCGCTGATGGAAGGCCAAGAATGGGACGTTTGGATCATGCAGACGAGCAGTCGTGGCC 794
          ||||||...|||||||||||||||||||||||||
Db   667 GGCGGCGTTCATCGGAAGGCCCTAGAAATCGGCCGGTTCGACCGAAGCCGGCCCATCGTAGCC 726

Qy   795 TTGCGAGGCCACAGCGCGTCCTTAGTAGTCCTCGGGAAACCTTCTTGCTTGGTTCGGTTGG 854
          ||||||||...||||||||||||||||||||||
Db   727 TTACGTGGACACAGTGCGTGCGTCCCTAGTTGTCTCCTCGTTCGTTTTCACTATGGTTCGGTTGG 786

Qy   855 TACGGATTTAACCCCGGTTCATTTAACAAAATCCTACTTACTTACGGTAACTCAGA--- 911
          | | |||||||||||||||||||||| ||||||| ||||||||||||||
Db   787 TACGGGTTCACCCCCGGTTCATTTCTCACCATAGCAAAAGCCTACGGTAACAACGGCGAA 846

Qy   912 ------AATTACACGGTCAATGGAGGCGCGGTTGGCAGAACCGCGGTCACCACTACCCTA 965
               ||||||||||||||||||| | |||||||||||||||||||
Db   847 AACGGTAATTACTACGGTCAGTGGAGGCGCGATAGGGCGCGATAGGGAGAAACAGCGGTCACAACCACATTG 906

Qy   966 GCGGGGTCAACAGCTGCCTTGACCACCGCTATTCGGTAAACGGGTGATATCCGGTCACTGG 1025
          | |||||||||||...||||||||||||||||||
Db   907 GCTGGATGCACTCGCGGCACTCCACCGCTTTTCAGCAAACGGTTATTGGAGGGCCACTGG 966

Qy   1026 AACGTGACCGATGTCTGCAACGGGCTGTTAGGCGGTTTCGCCGGCGATAACAGCCGGTTGC 1085
          |||||||||...||||||||||||||||||||||||||||
Db   967 AAAGTGCTTGACGTGTGTAACGGCCTGCTCGGCGGTTTCGCCGGCCATCACCTCCGGTTGC 1026

Qy   1086 TCCGTGGTTGGAGCCATGGACGCAGCCATCGTAATGCGGTTTGTTGCTTCTATAGTATTAATA 1145
          |||||||...||||||||||||||||||||||||
Db   1027 TCCGTTGCGTCGAGCCGTGGCCGGGCGATTGTCTGCGGTTTTGTGCGGTTGCAGCGTGGGTGCTGATC 1086

Qy   1146 GCTTTGCAACAAATTAGCAGAGAAGGTTAAGTTCGACGATCCTCTGGAGGCGGCGCAGTTG 1205
          ||||||||...||||||||||||||||||||||
Db   1087 GGGCTCCAATTTGGTTGCGGCGAAAATGAAGTCGACGATCCTCTGGAGGCGGCGCCAGCTT 1146

Qy   1206 CACGGTGGGTGTGGCACGTGGGGGGTGAATATTCACGGCGTTGTTCGCAAAAAGGAGTAT 1265
          |||||||||||||||...||||||||||||||||||||||
Db   1147 CACGGCGGGTGCGGCGCCGTGGGGGGTTTTATTCACGGGCGTGTTTGCGAAGGGGGAGTAC 1206

Qy   1266 GTGAAGGAGGTTACG---GGTTGGGGAGGGCCACGGGTTGCTCATGGGGGGTGGTGGG 1322
          ||||||||||||| ||||||||||||||||||||||||
Db   1207 GTGGGCGGAGGTTTACGGCAGCGCTGGGAGGCCGTAGGGGTTGTTAATGGGTGCGGAGGG 1266

Qy   1323 AAGTTGCTGCGGCGGCACGTGATTCAGATTCGGTGATTGCTGGGTGGGTAGTGCAGTTG 1382
          ||||||||...||||||||||||||||||||||
Db   1267 AAGTTGCTCGCGGCGGCGCAGTTGATTGAGATATTGGTGGTGTGTGGGGTTACGGCGACG 1326

Qy   1383 ATGGGACCCTGTTTTGGGGGTTGAATAAACTGAAGCTGTGAATAACACCATGAAC 1442
          |||||||...|||||||||||||||| ||||||
Db   1327 ATGGGGCCGGTTGTTCTCATGGGCTTCATAAGACGAAACTTTTGAGGATTTCTGAAGATGAT 1386

Qy   1443 GAGCTTGCGGGGGATGGACATGACTCGCCATGGAGGCTTTGCTTATGCTT------ATGAG 1496
          ||||||||...

```
Db    554 CCGGTCGTGTCCCACTGGATCTGGTCCGTCGGCTTGGCCTTCGCGGGCCCACGTCC  613
          |||||  |||||||||||||||| ||  |||||||||||||||||||||||||||
Qy    654 GACCGGCTATTTTCCA---CCGGCGTATAGACTTGGTCCGGTTCCGGTAGTCCACATG  710
          |||||  |||    ||||||||||| || ||||||||||||||||||||||||||||
Db    614 GGCCCGCTGCTCTTCAAGTCCGGCGTGATCGATCGACTTCGCCGGCTCCGGTCGCACATG  673

Qy    711 GTCGGCGGAATAGCCGGCCATGGGAGCGCTGATCGAAGGCCAAGAATGGACGTTTC  770
          ||||| |||| |||||||||||||||| |||||||| |||||   ||||||||||
Db    674 GTCGGGGCGATCGCCGGCTTCGTGGGGCGCGCTTCATCGAGGGCCCCATCGGCCGGTTC  733

Qy    771 GATCATGCAGGACGAGCTGTGCGCCTTGCGAGGCCACGCCGTCCTTAGTAGTCCTGGGA  830
          |||| |||  |   ||||||| |||||||| |||||  |||||||||||||||||||||
Db    734 GACCACGCCGGCCGTCGGGTGGCGCTCAAGGGCGCCCAAGCGCGTCGCTCGTCGTGGGG  793

Qy    831 ACCTTCTTGCTTGGTTGGTTACGGATTAACCCCGGTTCATTTAACAAAAATCCTA  890
          |||||| |||| |||  |||||||||| ||||||||||  |||| |||||  |||
Db    794 ACCTTCCGCTCGTGTTCGGCTGGTACGGGTTCAACCCGGGTTCCTTCGTCGTCCATCCTC  853

Qy    891 CTTACTTACGGTAACTCAGAAATTACTACGGTCAATGGAGAAATTACTACGGCC  950
          ||||  || ||  |   |||||||||| |||| ||||||||||||||| |||||
Db    854 AAGTCGTACGGCCCGCCCGGAGCATCAACGGCAGTGTCCGGCGTGGCCCGCACCGGCC  913

Qy    951 GTCACCACTACCCTAGCGGGTCAACAGCTGCCTTGACCACGCTATTCGGTAAACGGGTG  1010
          ||||||  |   || ||||||||||||| ||||||||||| ||  ||||||  |||||
Db    914 GTGACGACGACGCGTGCGGCGCAGCGTGGCGCGGGCTGTTAGGCGGTTTCGCGGG  973

Qy   1011 ATATCCGGTCACTGACCGATGTCTGCAACGGGCTGTTAGGCGGTTTCGCGGG  1070
          |||| |||| |   ||| |||| ||||  ||||||  ||||||||||||||||
Db    974 CAGACGGGGCACTGGAACGTGGTGGACGTCTGCAACGGCCTGCTCGGCGGGGTTCGCGGC  1033

Qy   1071 ATAACAGCCGGTGCTCCGTGGTTGAGCCATGGCAGCCATCGTATGCGGTTTGTTGCT  1130
          ||||||   || |  ||||||||||| ||  ||||||||||||| |||| |  |||||
Db   1034 ATCACCCGGGTGCAGCGGTGCGTGGTCGACCCGTGGCCCGTCATCTCGCGGCTTCGTCTCC  1093

Qy   1131 TCTATAGTATTAATAGCTTGCAACAAATTAGCAGAGAAGGTTAAGTCGACGATCCTCTG  1190
          |     ||| ||| |||| ||| ||  |||||||| |||||||||||||||||| |||||
Db   1094 GCCTGGGTGCTCATCGGGCTCAACGCGCTCAAGTACGACGACGACCCGCTG  1153

Qy   1191 GAGGCGGCGCAGTTGCACGGTGGGGTGCACGTGGGGGTGATATTCACGGCGTTGTTC  1250
          |||  ||| || |||||||| ||  ||||||||||||||| |  | |||| |||| |
Db   1154 GAGGCGGCAGCTGCACGGCGCTGCGGCGGTGGGGATCATCTTCACGGGCGCTGTTC  1213

Qy   1251 GCAAAAAGGAGTATGTGAAGGAGGTTAGGGGTTGGGAGGGCGCACGGGTTGCTCATG  1310
          || |||| | | |||||||| ||| || || |||||||||||| || |||||||||||
Db   1214 GCCAAGAAGCAGTACGTGGAGGAGATCTACGGCGCGGCCAGGCCGGTACGGGCGTTCCTG  1273

Qy   1311 GGGGGTGGGTGGAAGTTGCTGGCGGCGCACGTGATTCAGATTCTGGTGATTGCTGGGTGG  1370
          |||||| ||  ||| ||| |||| || |||||   ||  |||||  |||||||| |||
Db   1274 GGCGGCGGCGGCGAGCTGCTCGCGGCGCACATCGTGCGAGATCCTCGTCGATCTCGGCGGCTTC  1333

Qy   1371 GTTAGTGCGGACCATGGAGCCTTGTTTGTGGGGTTGAATAAACTGAAGCTGTTGAGAATT  1430
          || |||||| ||||||||  ||    ||||| |||||||||   |||| |||| ||||
Db   1334 GTGAGCTGCACCATGGACCCGCTCTTCTTGGCGCTCAAGAGCTGGGCCTGCTCCGCATC  1393

Qy   1431 TCTTCAGAGGATGAGCTTGCGCGGGATGGACATGACCTCGCCATGGAGGCTTTGCTTATGCT  1490
          |||||||| ||||||| ||| |  |||| |||| || ||||| ||| |||| ||||| |||
Db   1394 TCGCCGGAGGACGAGGTGGCCGGCATGAACGTGACCGACCGCACCGGTGGTTCGCCTACGTC  1453

Qy   1491 TATGAGGATGATGAGACGACGACAA  1513
          ||||||  |||  ||||||||||||
Db   1454 TACCACGACGACGACGAGCACGA  1476
```

RESULT 4
AX412659
LOCUS   AX412659   1506 bp   DNA   linear   PAT 14-JUN-2002
DEFINITION Sequence 423 from Patent WO0222675.
ACCESSION   AX412659
VERSION   AX412659.1 GI:21445117
KEYWORDS   .
SOURCE   Arabidopsis thaliana (thale cress)
  ORGANISM Arabidopsis thaliana

```
Db   1387 GAGACTGCGCGGGATGGATTGACGAGCATGGTGGGCATGCTTATGCTTACCATGATGAT  1446
          ||| |||||||||| ||||||| |||| ||||||||||||   ||||||||||||||||
Qy   1497 GATGATGAGACGCGACAAGCAGTTCGACGGGATGCAGTTGGGCCCAACGGCGTCTTCC  1556
          ||||||| |||| ||||||||||||||||||| ||||||||||||||||||||||||||
Db   1447 GATGATGTGTCCACCAAGAGAGGGGGTCATGATGAGTGAGTAGGCCTGGCTAGTAGTTCT  1506

Qy   1557 ACACCCACCA  1566
          ||  ||||||
Db   1507 CCCTCCACCA  1516
```

RESULT 3
BT009318
LOCUS   BT009318   1991 bp   mRNA   linear   PLN 20-JUN-2003
DEFINITION Triticum aestivum clone wlm12.pk0020.b10:fis, full insert mRNA
            sequence.
ACCESSION   BT009318
VERSION   BT009318.1 GI:32128869
KEYWORDS   FLI_CDNA.
SOURCE   Triticum aestivum (bread wheat)
  ORGANISM Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooideae; Triticeae; Triticum.
REFERENCE   1 (bases 1 to 1991)
  AUTHORS   Tingey,S.V., Wolters,P., Powell,W., Dolan,M., Miao,G.-H.,
            Caraher,N.R., Hanafey,M.K. and Hainey,C.F.
  TITLE   Direct Submission
  JOURNAL   Submitted (20-JUN-2003) Crop Genetics, E. I. DuPont de Nemours and
            Company, 1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104,
            USA
FEATURES   Location/Qualifiers
     source   1..1991
                     /organism="Triticum aestivum"
                     /mol_type="mRNA"
                     /db_xref="taxon:4565"
                     /clone="wlm12.pk0020.b10:fis"
ORIGIN

Query Match   45.1%; Score 768.6; DB 8; Length 1991;
  Best Local Similarity   73.8%; Pred. No. 2.9e-126;
  Matches 991; Conservative   0; Mismatches 349; Indels   3; Gaps   1;

```
Qy    174 GCGCCGCTGGACAGCAAGTTCGTCGCGACACGGCCTTCGCCGTCGACAACAACTACCTCCTC  233
          || | ||||| |||| |||||||||||  |||||||||||| ||||||| ||||||||||||
Db    134 GACTACCTGTCGAACACAGTTCGCGGCGTTCGCCGCACACAGGTTCGCCGGCTCCGGTCGTC  193

Qy    234 TTTTCCGCTACCCGTTTTTCATGCAGCTCAATGCAGCTCTCGCGCCGCCGGCTCC  293
          |||||||||||| |||  ||||  ||| |  |  ||| |||||||||||||||||||
Db    194 TTCTCGGCCTACCTCGTCTTCGGCGATCGAGCGTCGCCATGCGCTCGGCCGCCGGCTCC  253

Qy    294 GTCCGCGCCAAGAACACCATGAACACATCATGCTCACCAAGCTCCTGGACGCGTGCGCCGCCGGC  353
          |||||||| ||||||| |||| ||| ||||| ||||| ||||   |||||||||| || |||||
Db    254 GTCCGGGCCAAGAACACAGAACATCATCATGCTCACCAAGCGTCGACGCGTCGCACGCGCCGGC  313

Qy    354 GGCCTCCTCTACTACCTCTCGGCTTCCTTCGGCCTTTCGGCCTTCCCCCTCCAACGGCTTC  413
          |||||  | |||| |||| ||||||||||||||||||||||||||  |||||||||||||||
Db    314 GCGCTCTTCTACACCTCTCTCGGCCTTCGGCCTTCGGCCTTCGGGACGCGTCGAACGGCTTC  373

Qy    414 ATCGGTAAACATTTTCTTCGGCCTCAAGGACATCCCTTCATCCTTCATCGACTACAGCTAC  473
          |||||  | ||||||| ||||||||||||||||  | ||||||| |  |||||||| |||||
Db    374 ATCGGGAAGACACTTCTTCGGCCTCAAGGACATCATGCTCCTTCATCTCGCGAGACCGGCTTC  433

Qy    474 TTCCTTACCAATGGCCTTCGGCCTTCGGCCTATCTCATCTCCTTCCTCGGCTTCCTCCTCGGC  533
          |||||| ||| |||||||||||||||||||| ||| ||  ||| ||||||||||||||||||
Db    434 TTCTCTTTCCAGTGGCCTTCGGCCTTCGGCCATCATCTCATCCTCGGCTTCCTTCCCGGCTTC  493

Qy    534 GAACGCACACAGTTCGTGCCGTATCTCATCATCTCCTCCTTCCTCACCGGCTTCGTCTAT  593
          ||| || | ||||||| | |||||||| |||| ||||||||  ||| ||| ||||| |||
Db    494 GAGAGGACGCAGTTCGTCGTCGCGTATCTCATCATCTCCTTCTTCTTTACGGGGATTCGTCTAC  553

Qy    594 CCGGTCGTGTCCCACTGGATCTGGTCCCCAGACCGCGGCTGGCCTTCGCGGGCCTTTAAGATCACC  653
          |||||||||||||||||||| |||||||||| |||||||| |||||| |||||||  ||||| |||
```

```
Qy  913  ATTACTACGGTCAATGGAGCGGCGGTTGGCAGAACCGGGTCACCACTACCCTAGGCGGGT  972
         ||||| ||||| ||||||| || || ||||||||| |||||| ||| |||||||||||||
Db  821  CATACAACGGCCAGTGGAGCGCGGTCGACGACAGCTGTGATCACACAACGTTAGCTGGCT  880

Qy  973  CAACAGCTGCCTTGACCACGCTATTCGGTAAACGGGTAGATCCGTCACTGGAACGTGA   1032
         |||| |||||||||||| || ||  ||| ||||| |||||| ||||| ||||||||||
Db  881  GCACCGGCGGCGCTGACAACCCTATTTGGGAAACGTCTACTCTCGGGACATTCGAACGTCA  940

Qy  1033 CCGATGTCTGCAACGGGCGTGTTAGGCGGGTTCGCGGCGATAACAGCCGGTTGCTCCGTGG  1092
         |||||||||| ||||||||  || |||| ||||||||  || ||||||||| |||| || 
Db  941  CTGATGTATGCAACGGCCTCCTCGGAGGGTTTGCAGCCATAAACTGGTGGTGCTCTTGTCG  1000

Qy  1093 TTGAGCCATGGGCAGCCATCGTATGCGGTTTTGTTGCTTCTTCTATAGTAATAATAGCTTGCA  1152
         ||||| |||||||||||||||  |||| |||| || ||||||| |   ||||||  ||||||
Db  1001 TTGAGCCATGGGCTGCGATCATCTGCGGGTTCGTGCGGCGCCCTAGTCCTCCTCGATGCA  1060

Qy  1153 ACAAATTAGCAGAGAAGTTAAGTTCGACGATCCTCTGGAGGCGCGCAGTTGCACGGTG  1212
         ||||||||| | || |||  |||||||||||||||||| || || |||||||||  ||
Db  1061 ACAAGCTCGCTGAGAAGCTCAAATACGACGACCCCTTCTTGAGCGAGCACACTACACGGTG  1120

Qy  1213 GGTGTGCACGTGGGGGGGTGATATTCACGGCGGTTGTTCCGCAAAAAGGAGTATGTGAAGG  1272
         ||||  ||||| |||  |     || ||||| ||| || || ||||| || || ||||||
Db  1121 GTTGCGGTGCGTGGGGACTAATATTCACGGCTCTCTTCGCTCAAGAAAGTACTTGAACC  1180

Qy  1273 AGGTTTACGGTTG------GGGAGGGGCCACGGGTTGCTCATGGGGGGGTGTGGGAAGT  1326
         |||| |  |||| |      ||  || |||| |||| ||||  ||||| ||||||  ||
Db  1181 AGATTTACGGCAACAAACCCGGAGGGCCACACACGTTTGTTTATGGGCGGTGGAGAAAAC  1240

Qy  1327 TGCTGGCGGCCACGTGATTCAGATTCTGGTGATTGCTGGGTGGGTTAGTGCGCACCATGG  1386
         || |||||  | |  ||| ||||||| |||| || ||| |||||||  ||||||||||||
Db  1241 TACTTGGAGTCAGCTGATCATTGTGATCATTGTGATCACCGGTTGGGTAAGTGCGCACCATGG  1300

Qy  1387 GACCCCTGTTTTGGGGTTGAATAAACTGAAGCTGTTGAGGATTTCTTCAGAGGATGAGC  1446
         ||||   ||||  ||||  || |||| || |||  ||| ||||||||||||||||||||
Db  1301 GGACACTTTTCTTCATCCTCAAGAAAATGAAATTGTTGCGGATATCGTCCGAGGATGAGA  1360

Qy  1447 TTGCGGGGATGGACATGACTCGCCATGGAGCCTTGCTTATGCTTATGAGGATGATGATGAGA  1506
         || |||  ||| | ||| | || |||  |||||||| ||| || |||| |||||||||||
Db  1361 TGGCCCGTATGGATATGACCACCAGGCACGCGGTGGTTTTGCTTATATGTACTTTGATGATGATG  1420

Qy  1507 CGCACAAGCAATGGGATGCAGTTGAGGAGGGTTGGGCCCAACGCGTCTTCCCACACCCACCA  1566
         ||||||||| || || ||  |  || ||||| ||||| || ||||||  |||| | ||||
Db  1421 AGTCTCACAAAGCCATTCAGCTTAGGAGAGTTGAGCCACGATCTCCTTCCTCTCCTTCTGGTG  1480

Qy  1567 CTGAT  1571
         || ||
Db  1481 CTAAT  1485
```

```
RESULT 5
AX506508
LOCUS       AX506508              1506 bp    DNA    linear   PAT 27-SEP-2002
DEFINITION  Sequence 1203 from Patent WO0216655.
ACCESSION   AX506508
VERSION     AX506508.1  GI:23387745
KEYWORDS    .
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1
  AUTHORS   Harper,J.F., Kreps,J., Wang,X. and Zhu,T.
  TITLE     Stress-regulated genes of plants, transgenic plants containing
            same, and methods of use
  JOURNAL   Patent: WO 0216655-A 1203 28-FEB-2002;
            The Scripps Research Institute (US) ; Syngenta Participations AG
            (CH)
FEATURES             Location/Qualifiers
     source          1..1506
                     /organism="Arabidopsis thaliana"
                     /mol_type="unassigned DNA"
```

```
REFERENCE   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
         1
  AUTHORS   Glazebrook,J., Wang,X., Dangl,J.L., Eulgem,T. and Zhu,T.
  TITLE     Plant genes, the expression of which are altered by pathogen
            infection
  JOURNAL   Patent: WO 0222675-A 423 21-MAR-2002;
            Syngenta Participations AG (CH) ; UNIVERSITY OF NORTH CAROLINA AT
            CHAPEL HILL (US) ; Glazebrook, Jan (US) ; Wang, Xun (US) ; Dangl,
            Jeffrey L. (US) ; Eulgem, Thomas (US)
FEATURES             Location/Qualifiers
     source          1..1506
                     /organism="Arabidopsis thaliana"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:3702"
ORIGIN
```

```
Query Match    44.3%;   Score 755.8;  DB 6;  Length 1506;
Best Local Similarity    71.7%;   Pred. No. 5.4e-124;
Matches 1036;  Conservative    0;  Mismatches  397;  Indels   12;  Gaps    3;

Qy  136  ACGGCTCCGCCGCCGCCTCCTTATCTGCGGCCGCCATTCGCCGGCCGTGGACAGCAAGTTCG  195
         |||  ||||| |||  ||| ||||||| ||  |||||||||||  || |||||||| |||
Db  44   ATGCCACGGCGGCGGCCAACTACATCTGTGGCCAGTTAGGCGACGTCAACAACAAATTA  103

Qy  196  TCGACACGGCCTTGCGTGACAACACCTACCTCCTTTTCGGCCTACCTCGTTTTT  255
         |||||||  ||| |||||  | |||| |||| ||| || |||| |||||||| |
Db  104  TCGACACCCGCTTTGCCTTTCGCTATAGACACACCTTACCTTCTTCTCCGCCTACCTTGTCTTCT  163

Qy  256  CTATGCAGCTCGGCTTCGCCATGCTCTGCGCGGCTCCGTCCGCCAAGAACACCACCATGA  315
         ||||| |||||||| |||  ||||||||||||| ||||| ||  ||| ||||| | ||||
Db  164  CTATGCAGCTGGGCTTCGGCTATGCTCGTGTGCCGGGTTCCGGTTCCGTGAAGAGCCAAGAATACTATGA  223

Qy  316  ACATCATGCTCACCAAGTCTCGGACGCTGCCGACGCGGCCGCCTCTTCTACACTCCTTCG  375
         || ||||||  ||||||| || ||||| | || || || |||| || ||| |||| |||
Db  224  ACATCATGTTCACCAAGTCTTGACGCTGCGAGCGCCGGTCCTCTTCTATTATCTGTTTG  283

Qy  376  GCTTCGCCTTCGCTTCGGCTCCCCCTCCAAACGCGTTCATCGTAAACATTTCTTCGGCC  435
         ||||||||||| |||  ||| |||| || |||| || |||| |  ||||| ||||||||
Db  284  GCTACGCCTTTGCCTTTGGATCTCCCAATGGTTTCATCGGTAAACATCACTTTTGGTC  343

Qy  436  TCAAGGACATCCCTTCATCCTCCTACGACTACTTCCTCTACCAATGGGCCTTCG  495
         |||||||||||||  | || ||| |||||||||||||| |||||||||||||||
Db  344  TCAAAGACATCCCACGGGCCTCTGCTGATCACTACTCCAACTTTCTCTACCAATGGGCCTTTG  403

Qy  496  CCATGCGCGCCGGCGGGCGGGCGATCACCAGGCGGAAGCATCGCCGAACGCACACAGTTCGTGCCT  555
         |||||||  ||| ||||||| || ||||| || ||||| | ||| ||||||||| ||||||||||
Db  404  CAATCGCTGCGGCGGAATCACACAAGTGGCGTCGATCGCTCGAACGGACCAGCAGTTCGTGCGCTT  463

Qy  556  ATCTCATCATCACTCCTCCTTCCTCACCGGGCTTCGTCTATCGGTGGGTCTCCCACTGGTTCT  615
         ||| |||||| | || ||| |||| | |||||||||| || ||||||||||| ||| || ||
Db  464  ACCTAATCTATTCCTCTTTTCTTAACCGGGTTTGTTTTTACCCGGTCGTCTCTCACTGGTTCT  523

Qy  616  GGTCCCCAGACGCGGTGGGCTGGGCCTCTGCCTT---TAAGATCACCGACCGGCCTATTTCCACCG  672
         ||||   ||| |  || || ||| ||| |||||       |||| ||||||  ||||| |||| |
Db  524  GGTCAGTTGATGGATGGGGCCAGCCCGTTCCGTACCGATGGAGATTTGCTTTTCAGCACCG  583

Qy  673  GGGTAATAGAGACTTCGCCCGGTTCCGGCGTAGTCCACATGGTCGCGGGAATAGCCGGCCTAT  732
         ||||| || || || ||| ||||  ||||  || ||  ||||| ||||| ||| ||||||||
Db  584  GAGCGATAGATTTCGCTGGCGGTTGTGTTGTTTCATATGGTCGGAGTATCGCTGGACTCT  643

Qy  733  GGGGAGCGGCTGATCGAAGGCCCAAGAAATGGGACGTTCGATCATCGCAGGACGAGCTGTGG  792
         ||||| | ||  |||| || ||||| | |||| ||||| ||||| || ||| || ||| ||
Db  644  GGGGTGCCGCTCATCGAAGGTCCACGGACTTGCCCGGTTCGATAACCGAGGCCGTGCCATCG  703

Qy  793  CCTTGCGAGGCCACAGCGCGGCGTCCTAGTAGTGCTTCTTGTTTGGTTCGGTT  852
         ||||  |||  ||||||||  ||| ||||| |||  || ||||| |||| |||
Db  704  CTCTTCGGTGGCCACTCGGCGTCACTTGTTGTCCTTCGGAACCATTCCTCTCCTTGGTTTGGAT  763

Qy  853  GGTACGGATTTAACCCCGGTTCATTTAACAAAATCCTACTTACTTACGGTAACTCAGGAA  912
         ||||||||||||| ||||||||| ||  | | |||| |||| ||  ||||||||||||||
Db  764  GGTACGGATTTAACCCCGGTTCCTTCAACAAGATCCTAGTAGTCAGCTACGCACGACGAGA--CAGGCA  820
```

```
ORIGIN          /db_xref="taxon:3702"

Query Match        44.3%;   Score 755.8;  DB 6;  Length 1506;
Best Local Similarity   71.7%;   Pred. No. 5.4e-124;
Matches 1036;  Conservative 0;  Mismatches 397;  Indels 12;  Gaps 3;

Qy    136  ACGCCTCCGCCGCGCCTCCTTATCTGCGGCCATTCGCCGCCGTGGACAGCAAGTTCG  195
           |||||||||| |||| ||| ||||| ||| ||||| |  || |||||  ||||||||
Db     44  ATGCCACGGGCGGCGGCAACTCACATCTGTGGCCAACTCAACAACAAATTA      103

Qy    196  TCGACACGGCCTTCGCCGTCGACAACACCTACTTCGCCTCCTTTTCCGCTACCTCGTTTTT  255
           | |||||||||||||||||||||||   |||||| ||| |||||||||||||||| ||
Db    104  TCGACACCGCTTTCGCTATGAGACAACACTTACCTTCTCTCCGCCTACCTTGTCTTCT     163

Qy    256  CTATGCAGCTCGGCGTTCGCCATGCTCTGCGCGGCTCCGTCCGCCGAAGACACCACCATGG  315
           ||||||||||| ||||| ||| || ||||||||||||||||||||||| ||| |||||||
Db    164  CTATGCAGCTTGGCTTCGGCTATGCTCTGCGTGTCCCGTTGAGAGCCAAGAGTACTATGA  223

Qy    316  ACATCATGCTCACCAAACGTCCTGGACGCGTCCCGCGGCGGCCTCTTCTACTACCTCTTCG  375
           ||||||||||| |||||| |||||| ||||| |||||||| ||||  | |||||| ||||
Db    224  ACATCATGCTTACCAAACGTCCTTGACGCGTGGTCTCTGTCTCTTATATCTGTTTG      283

Qy    376  GCTTCGCCTTCGCTTCGGCTCCCCCTCCAACGGCTTCATCGTAAACATTCTTCGGCC     435
           |||||||||| ||| ||| ||| ||||||||||| |   |||||||||| ||||| ||
Db    284  GCTACGCCCTTGCCTTGATCTCCGTCCAATGGTTTCATCGGTAAACACTACTTTGGTC     343

Qy    436  TCAAGGACATCCCTCATCCTCCTACGACTACAGCTACTTCCTCTACCAATGGGCCTTCG    495
           ||||  ||||| || ||| |||| |  ||||||||| |  |||||  || |||||||
Db    344  TCAAAGACATCCCCACGGCCTCTGCTGACTACTCCAACTTTCTCTACCCAATGGGCCTTTG  403

Qy    496  CCATCGCGCCGCCGCCGGCATCACCAGCGGAAGCATCGCGAACGCACAGTTCGTGGCCT    555
           ||||||| | ||| |  || ||||||||| |||  |||| |||||||||||||| ||||
Db    404  CAATCGCTGCGGCTGGAATCACAAGTGGCTCGACGCTGAACGGAGCACAGTTCGTGGCTT  463

Qy    556  ATCTCATCTCACTACTCCTCCTTCGTCATACGCGGTCTCTCCCACTGGTTCT         615
           |||| || || |  |||||| |||||||  ||||| |||||| |||||||||
Db    464  ACCTAATCTATTCCTTCTTTCTTAACCGGTTGTTTACCCGTGTCTCTCCACTGGTTCT     523

Qy    616  GGTCCCCAGACGCGTGGGCCTCTGCCTT--TAAGATCACCGACCGGCTATTTTCCACCG    672
           |||||           || ||||  | |       |   ||| || | |||||||||||
Db    524  GGTCAGTTGATGGGATGGGCGCAGCCGGCCGTTCGCTTTTGCTTTTCAGCACCG      583

Qy    673  GGGTAATAGACTTCGCCGGTTCCGGCGTAGTCCACATGGTCGGCGGAATAGCCGGCCTAT  732
           |   || |||| |||| ||  ||| |  | ||  |||| |||| |  | || ||| ||
Db    584  GAGCGGATAGATTTCGTCGGTGTGTTCATATGGTTCGGCGTATCCTGTGACTCT      643

Qy    733  GGGGAGCCGCTGATCGAAGGCCAAGAAATGGGACGTTTCGATCATCGAGGACGAGCTGTGG  792
           |||| |||| || || |   || ||| |||||  |   |||  |||| ||||| |||| |
Db    644  GGGGTGCCGTCATCGAAGGTCCACGACTTGGCCGGTTCGATAACCGAGGCCGTGCCATCG  703

Qy    793  CCTTGGCGAGGCCACAGCGGCGCTCCTTAGTAGTCGTGGAACCTTCTTGCTTTGGTTCGGTT  852
           ||  ||||||| || |  |  ||||| |||| |  |||| |||||| ||| |||||| ||
Db    704  CTCTTCGTGGCCACTCGGCGTCACTGTTCCTTGGAACATTCCTCTCTGGTTGGATTTTTGGAT  763

Qy    853  GGTACGGATTTAACCCGGTCATTCATTAACAAAATCCTACTTACTACGGTAACTCAGGAA   912
           |||||| ||| |||||| | | |||| || |||| ||  || || ||||||| ||||
Db    764  GGTACGGATTTTAACCCGGTTCCTTCAACAAGATCCTACTAGTTCACGTCACGTACGAGA--CAGGCA  820

Qy    913  ATTACTACGGTCAATGAGCGCGGTTGGCAGAACCGCGGTCACCACTACCCTAGCCGGGGT   972
           |||| | || || || |  |  ||||||||| ||||| | ||||||||| || ||||| |
Db    821  CATACAACGGCCAGTGAGCGGACGGGCGGAGCAGTGTGCACAACAACGTTAGCTGGCT   880

Qy    973  CAACAGCTGCCTTGACCACGCTATTCGGTAAACGGTGATATCCGTCACTGGAACGTGA     1032
           || ||| || || ||||||||||  |||| |||||||| |||  |||||||||||||||
Db    881  GCACCGCGCGGCTGACAACCCTATTTGGAAACGTCTACTCCTCGGGACAATTGGAACGTCA  940

Qy   1033  CCGATGTCTGCAACGGGCTGTTAGGCGGTTTCGCGGCGATAAACAGCGCGGTTGCTCCGTGG  1092
           | ||||  |||||||| |    |    | |  || ||| ||| ||| | | |||||| ||
Db    941  CTGATGTATGCAACGGGCGCCTCCTCCTCGGAGAGGTTTGCGCCATAACTGGTGGCTCTCGTCG  1000

Qy   1093  TTGAGCCATGGGCAGCCATCGTATGCGGTTTTGTTGCTTCTATAGTATTAATAGCTTGCA  1152
           |||||| |||| | | |  |
Db   1001  TTGAGCCATGGGCTGCGGATCATCTCGCGGGTTGTGTGGCGCCCCTAGTCCTCCTCGGATGCA  1060

Qy   1153  ACAAATTAGCAGAGAAGGTTAAGTTCGACGATCCTCTGGAGGCGGCGCAGTTGCACGGTG  1212
           ||||| ||    |||||||||||| ||||||||   ||||||| ||  |  ||  | |||
Db   1061  ACAAGCTCGCTGAGAAGCTCAAATACGACGACCCTCTGAGCGAGCACAACTACACGGTG  1120

Qy   1213  GGTGTGCGCACGTGGGGGTGGTGATATTCACGGCGGTGTTCGTCGCAAAAAGGAGTATGTGAAGG  1272
           ||   || |   |||
Db   1121  GTTGCGGTGCGTGCGGGGGACTAATATTCACGGCTCTCTTCGCTCTCAAGAAAGTACTTGAACC  1180

Qy   1273  AGGTTTACGGTTG-----GGGAGGGCGCACCGGTTGCTCATCGGGGTGGTGGGAAGT  1326
           |   |||| |||       |||||||||| || |     | ||  |||| |||||
Db   1181  AGATTTACGGCAACAAACCCGGAAGGCCACACGGTTTGTTTATGGGCCGGTGGAGGAAAAC  1240

Qy   1327  TGCTGGCGGCGCACGTGATTCAGATTCGGTGATTGCTGGGTGGGTGGGTTAGTGCGACCATGG  1386
           || ||||| |  || || ||||||||   |||  |  ||| || ||| |||  | ||||||||
Db   1241  TACTTGGAGCTCAGCTGATTCAGATCATTGTGATCACCGGGTTTGGGTAAGTGCGACCATGG  1300

Qy   1387  GACCCTGTTTTGGGGTTGAATAAACTGAAGCTGTTGAGGATTCTTCAGAGGATGAGC  1446
           |||| | ||| |   |||  | ||| |  || |  |||| |||| ||||||| || |
Db   1301  GGACACTTTCTTCATCCTCAAGAAAATGAAATTGTTGCGGATCGTCCGGAGGATGAGA  1360

Qy   1447  TTGCGGGGATGGACATGACTCGCCATGGAGGCCTTTGCTTATGCTTATGAGGATGATGAGA  1506
           || |   | ||||| ||| ||| ||||| |||| |  |  | | ||| |||||| ||||
Db   1361  TGGCCGGTATGGATATGACCAGGCACGGTGGTTTTGCTTATATGTTGATGATGATG  1420

Qy   1507  CGCACAAGCACGGATGCAGTTGAGGAGGGGTTGGGCCCAACGGCGTCTTCCACACCCACCA  1566
           |  || ||||   ||||| |||||
Db   1421  AGTCTCACAAAGCCATTCAGCTTAGGAGGAGTTGAGCCGGTTAGGAGAGTTGAGCCTCTGGTG  1480

Qy   1567  CTGAT      1571
           || ||
Db   1481  CTAAT      1485


RESULT 6

ATAMT1

LOCUS       ATAMT1                  1700 bp    mRNA    linear   PLN 11-JUN-1997
DEFINITION  A.thaliana amt1 mRNA.
ACCESSION   X75879
VERSION     X75879.1  GI:551218
KEYWORDS    amt1 gene.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1
  AUTHORS   Ninnemann,O., Jauniaux,J.C. and Frommer,W.B.
  TITLE     Identification of a high affinity NH4+ transporter from plants
  JOURNAL   EMBO J. 13 (15), 3464-3471 (1994)
  MEDLINE   94341258
   PUBMED   8062823
REFERENCE   2  (bases 1 to 1700)
  AUTHORS   Frommer,W.B.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-JAN-1994) W.B. Frommer, Institut fuer Genbiologische,
            Forschung Berlin GmbH, Ihnestr 63, 14195 Berlin, FRG
FEATURES             Location/Qualifiers
     source          1..1700
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /db_xref="taxon:3702"
                     /tissue_type="complete seedling"
                     /dev_stage="seedling"
                     /note="ecotype Landsberg"
     gene            1..1700
                     /gene="amt1"
     CDS             21..1526
                     /gene="amt1"
                     /codon_start=1
                     /protein_id="CAA53473.1"
                     /db_xref="GI:551219"
```

```
/db_xref="GOA:P54144"
/db_xref="SWISS-PROT:P54144"
/translation="MSCSATDLAVLLGPNATAAANYICGQLGDVNNKFIDTAFAIDNT
YLLFSAYLVFSMQLGFAMLCAGSVRAKNTMNIMLTNVLDAAAGGLFYLFGYAFAFGS
PSNGFIGKYHYFGLKDIPTASADYSNFLYQWAFAIAAAGITSGSIAERTQFVAYLIYSS
PLTGFVYPVVSHWFWSVDGWASPFRTDGDLLFSTGAIDFAGSGVVHMVGGIAGLWGAL
IEGPRLGRFDNGGRAIALRGHSASLVVLGTFLLWFGWYGFNPGSFNKILVTYETGTYN
GQWSAVGRTAVTTLAGTAALTILFGKRLLSGHWNVTDVCNGLLGGFAAITGGCSVV
EPWAAIIGFVAALVLLGCNKLAEKLKYDDPLEAAQLHGGGAWGLIFTALFAQEKYL
NQIYGNKPGRPHGLFMGGGGKLLGAQLIQIIVITGWVSATMGTLFFLIKKMLLRISS
EDEMAGMDMTRHGFAYMYFDDDESHKAIQLRRVEPRSPSPSGANTTPTPV"
```

ORIGIN

```
Query Match      44.3%;  Score 755.8;  DB 8;  Length 1700;
Best Local Similarity   71.7%;  Pred. No. 5.4e-124;
Matches 1036;  Conservative   0;  Mismatches 397;  Indels 12;  Gaps   3;
```

```
Qy   136 AGGCTCCGCGCGGCCTCCCTTATCGTGCGGCCATTTCGCCGTGGACAGCAGCAAGTTCG  195
             ||| ||||||| ||| |||||||| | |||| ||||||||||||| |||||  || ||
Db    64 ATGCCACGGGCGGCGGCCAACTACATATGTGGCCAGCTGACGTCAACACAAATTCA    123

Qy   196 TCGACACGGCCTTCGCCGTGCACACACCTACCTCCTCTTTTCGCCTACCTCGTTTTTT  255
             |||||||||||||||| |||  |  || |||||||| |||||||  ||||||||||
Db   124 TCGACACCGGTTCGCTATAGACAACACTTACCTCCTCTTCTCCGCCTACCTTGTCTTCT   183

Qy   256 CTATGCAGCTCGGCTTCGCCATGCTCTGCGCCGGCTCCGTCCGCGCCAAGAACACCATGA  315
             ||||||||||||||||| |||||||||||| ||||| ||  | ||||| || ||
Db   184 CTATGCAGCTTGGCTTCGCTATGCTCTGTGCCGGTTCGGTAGAGACAACTATGA        243

Qy   316 ACATCATGCTCACCAACGTCCTGGACGCTGCCGGCGGCCTTCTACTACCTCTTCG  375
             ||||||||||||||||| ||||||||| ||||| |||||  | ||||||||||||||
Db   244 ACATCATGCTTACCAACGTCCTTGACGCTGCAGCCGGTGTCTTCTTCTTATATCGTTTG   303

Qy   376 GCTTCGCCTTCGCTTCGCCTCCCCTCCAACGGCTTCATCGTAAACATTTCTTCGGCC  435
             ||||||||||||| || ||||||| |||| |||||||||| ||||||| ||||||| ||
Db   304 GCTACGCCTTTGCCTTTGCCTCCGTCCCAATGGTTTCATCGTAAACACTACTTTGGTC   363

Qy   436 TCAAGGACATCCCTTCATCCTCCACGACTACAGCTACTTCCTCTACCAAATGGGCCTTCG  495
             |||| ||||||||||||||| ||| ||||||||||  | |||||||||||||||| ||
Db   364 TCAAAGACATCCCCACGGCCTCTGCTGACTACTCCAACTTTCTCTACCAAATGGGCCTTTG  423

Qy   496 CCATCGCCGCCGGCCATCACCAGCGGAGCATCGCCAGACACGCACACAGTTCGTGCGCT  555
             ||||| | |||| || ||||||||||||||||| | | || ||||||||||| |||| ||
Db   424 CAATCGCTGCGGCTGGAATCACAAGTGGCTCGATCGCCAGCGGACACCAGTTCGTGGCTT  483

Qy   556 ATCTCATCTACTCCTCCTTCGTCTATCCGGCGCTTCTCATCCGGTGGTCTCCCCACTGGTTCT  615
             ||| |||| |  ||| |||||| ||||| || || |||||||| |||| || || ||||
Db   484 ACCTTAATCTATTCCTTTCTTTCTTAACCGGGTTTGTTTACCGGTCGTCTCTCACTGGTTCT  543

Qy   616 GGTCCCCAGCAGCGGCTGGGCCTTGCCTT---TAAGATCACCGACCGGCTATTTTCCACCG  672
             ||  ||  |||||||||||| |||||||||     || | |||||||||||| |||  ||
Db   544 GGTCAGTTGATGATGGGGCCAGCCCGTTCCGTAGCCGATGGAGATTTGCTTTTCAGCACCG  603

Qy   673 GGGTAAATAGACTTCGCCGGTTCCCGGCGTAGTCCACATGGTCGGCGAATAGCCGGCCTAT  732
             ||| | || |||| ||| ||  || ||||| ||||| ||||| |||||| ||  |||||
Db   604 GAGCGATAGATTTCGCCGGTCCCGGTGTTCATAATGGTCGGAGTATCCGCTGGACTCT     663

Qy   733 GGGGAGCGGCTGATCGAAGGCCAAGAAATGGAGACGTTGATCATGCGAGGACGAGCGTGG  792
             ||||| ||| || |||||||| |||  |||||| |||| || ||||| | |||||||||
Db   664 GGGGTGCGGTCATCGAAGGTCCACGACTTGGCCGGTTCGATAACCGAGGCCGTGCCATCG  723

Qy   793 CCTTGCGAGGCCACAGCGGCGTCGTTGGTCCTTGGTTCGGTT  852
             |||||| ||||| ||| || ||| ||| || ||||| |||| || |||| ||
Db   724 CTCTTCGTGGCCACTCGGCGTCACTTGTTGTCCTTGGAACATTCCCTCCTCGGTTTTGGAT  783

Qy   853 GGTACGGATTAACCCCGTTCATTTAACAAAATCTACTTACTTACGGTAACTCAGGAA  912
             ||| ||||||||||||||||| ||||  ||||||| | |||||  |||||| |||||||
Db   784 GGTACGGATTTAACCCCGGTTCCTTCAACAAGATCCTAGTCACGTACGAGA--CAGGCA    840

Qy   913 ATTACTACGGTCAATGAGCGCGGTTGGCGAGAACCGCGTCACCACTACCCTAGCGGGT  972
             |||| | || || | |  | ||||||| |||| || | || |||||  | |||||||||
Db   841 CATCACACGGCCAAGTGACAGCGTCGGACGGCAGCGTCACAACAAGTGTCACAGGTGGCT  900
```

```
Qy   973 CAACAGCTGCCTTGACCACGCTATTCGGTAACGGCGTAAACCGGTGATCATCCGTCACTGGAACGTGA  1032
             | |||||| ||  || || ||| | |||| || |||| ||||||  ||||| || || ||||||
Db   901 GCACCGGCGGCGCTGACAACCCATTTGGAAACAGTCTACTCTCGGGGAACATTGGAACGTCA      960

Qy  1033 CCGATGTCTGCAACGGCTGTTAGGCGGTTCGCCGGGGGATAAACAGCCGGTTGCTCCGTGG  1092
             |  |||| || ||||||| |||||||| |||| ||||||||| || |||||| | |  ||
Db   961 CTGATGATGCAACGGCCTCCTCCGGAGGGTTGCAGCCGAGGGTTGCTGCTCGTGG   1020

Qy  1093 TTGAGCCATGGGCAGCCATCGTATGCGCGGTTCGGTTTGTTGCTTCTATAGTATTAATAGCTTGCA  1152
             ||||||||||||||||||||| || ||||| ||||||||||||||||||| |||||| ||||
Db  1021 TTGAGCCATGGGTGCGATCATCTGCGGGTTCGTGCGGGGCCCTAGTCCTCCTCCGGATGCA   1080

Qy  1153 ACAAATAGCAGGAAGGTTAAGTTCGACGATCCTCCTGAGGCGGCGGCAGTTGCACGGTG  1212
             |||||||| | ||| ||| | |||| |  | ||| |||||||||||||| ||||| |||||
Db  1081 ACAAGCTCGCTGAGAAGTCAAATACGAGCACCCTCTTGAGGCAGCACAACTACACGGTG   1140

Qy  1213 GGTGTGGCACGTGGGGGGGTGATATTCACGGCGTTGTTCCGCAAAAAGGAGTATGTGAAGG  1272
             ||  || |||||| | ||||| ||||||| |||| ||  | |||||||| ||| |||| |||
Db  1141 GTTGTCCGGTGCGTGGGAGCTAAATATTCACGGCGTCTCTTGCTCTTCGCTCTCAAGAAAGTACTTGAACC   1200

Qy  1273 AGGTTTACGGGTTG------GGGAGGGGCGCACGGGTTGCTCATGGGGGTGGTGGGAAGT  1326
             |||||| || |||        || || |||  | || ||| ||| || | ||| |||||
Db  1201 AGATTTACGGCAACAAACCCGGAGGCCACACCGGTTTGTTTATGGCGGTGGAGGAAAAC   1260

Qy  1327 TGCTGCGGCGGCACGTGATTCAGATTCTCGTGTATGATCGTGGTGGGTGGTAGTGCGACCATGG  1386
             || | |||| | |||| ||||||| | |||| ||||| ||| | ||| |||| ||||||
Db  1261 TACTTGGAGCTCAGCTGATTCAGATCATTGTGATCACCGGTTGGGTAAGTGCGACCATGG    1320

Qy  1387 GACCCTTGTTTTGGGGGTGAATAAAACTGAAGCTGTTGAGGATGAGC  1446
             |||| |  | | |||||| |||| ||  |  | |||| |||||| | |||||||||
Db  1321 GGACACTTTTCTTCATCCTCAAGAAATGAAATGTTGCGGAATATCGTCCGAGGATGAGA    1380

Qy  1447 TTGCGGGATGACATGACTCGCCATGGAGGCTTTGCTTATGCTTATGATACTTTGATGATGATGA  1506
             |||| ||||| ||||| || || ||| |||  |||  | ||| ||||  |  ||||||| |||
Db  1381 TGGCCGGTATGGATATGACCAGGCACGGTGGTTTGCTTATATGTACTTTGATGATGATG   1440

Qy  1507 CGCACAAGCATGGGATGCCAGTTGAGGAGGGGTCTTCCACACCCACCA  1566
             || |||| ||| | |||| || || |||| |||  ||||||||| | |||||
Db  1441 AGTCTCACAAAGCCATTCAGCTTAGGAGAGTTGAGCCGCACGATCCTCCTTCTCCTTCTGGTG    1500

Qy  1567 CTGAT  1571
             |||||
Db  1501 CTAAT  1505
```

RESULT 7
A44314
LOCUS   A44314   1748 bp   DNA   linear   PAT 07-MAR-1997
DEFINITION Sequence 1 from Patent WO9511978.
ACCESSION   A44314
VERSION   A44314.1 GI:2299142
KEYWORDS   .
SOURCE   Arabidopsis thaliana (thale cress)
ORGANISM   Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1 (bases 1 to 1748)
AUTHORS   Frommer,W. and Ninnemann,O.
TITLE   DNA SEQUENCES FOR AMMONIUM TRANSPORTER, PLASMIDS, BACTERIA, YEASTS,
            PLANT CELLS AND PLANTS CONTAINING THE TRANSPORTER
JOURNAL   Patent: WO 9511978-A 1 04-MAY-1995;
            INST GENBIOLOGISCHE FORSCHUNG (DE)
COMMENT   Other publication CA 2175211 950504
            Other publication AU 8060094 950522
            Other publication DE 4337597 950504.
FEATURES   Location/Qualifiers
     source   1..1748
                     /organism="Arabidopsis thaliana"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:3702"
     CDS   21..1526
                     /function="TRANSPORT OF AMMONIUM ACROSS MEMBRANES"

/codon_start=1
/product="AMMONIUMCARRIER"
/protein_id="CAA02780.1"
/db_xref="GI:2299143"
/db_xref="REMTREMBL:CAA02780"
/translation="MSCSATDLAVLLGPNATAAANYICGQLGDVNNKFIDTAPAIDNT
YLLFSAVLVFSMQLGFAMLCAGSVRAKNTMNIMLTNVLDAAAGGLFYYLFGYAFAFGS
PSNGFIGKHYFGLKDIPTASADYSNFLYQWAFAIAAGITSGGSIAERTQFVALLIYSS
FLTGFVYPVVSHWFWSVDGWASPFRTDGDLLFSTGAIDFAGSGVVHMVGGIAGLWGAL
IEGPRLGRFPDNGRAIALRGHSASLVVLGTFLWFGWYGFNPGGSFNKILVTYETGTYN
GQWSAVGRTAVTTLLAGCTAALTTLFGKRLLSGHWNVTDVCNLLGGAWGLIFTALFAQ
EPWAAIICGFVAALVLLGCNKLAEKLKYDDPLEAAQLHGGCGAWGLIFTALFAQBKYL
NQIYGNKPGRPHGLEMGGGKLLGAQLQIIVVTGWVSATMGTLFFILKKMKLRISS
EDEMAGMDMTRHGGFAYMYFDDDESHKAIQLRRVEPRSPSPSGANTTPTPV"
ORIGIN

Query Match   44.3%;   Score 755.8; DB 6;   Length 1748;
Best Local Similarity 71.7%; Pred. No. 5.4e-124;
Matches 1036; Conservative   0; Mismatches 397; Indels 12; Gaps   3;

Qy   136 ACGCCTCCGCCGCCGCCTCCCTTATCTGGGGCCATTGCCCGCCTGGACAGCAAGTTCG 195
            |||| | ||||| || || |||||||| || | |||| |||||| | ||||
Db   64 ATGCCACGGCGGCGGCCAACTACATATGTGGCCAGCTAGGCGACGTCAACAACAAATTCA 123

Qy   196 TCGACACGGCCTTGCCGTGCAGACAACCTACCTCCTTTTCCGCCTACCTCGTTTTT 255
            |||||| |||||| ||| | ||| ||||| || ||| | || ||||| |||||||
Db   124 TCGACACCGCCTTTGCGTATAGACAACACTTACCTCCTCTCCGCCTACCTTGTCTTCT 183

Qy   256 CTATGCAGCTCGGCTTCGCCATGCTCTGCGCCGGCTCCGTCCCGGGCCAAGAACACCATGA 315
            ||||| | ||| |||||| ||||| ||| || || | |||| || |||||| |||||
Db   184 CTATGCAGCTTGGCTTCGCTATGCTCTGTGCCGGTTCCGTCCGTGCGAGAGCCAAGAATACTATGA 243

Qy   316 ACATCATGCTCACCAACGTCCTGGACGCGTCGCCGCCGGCGCCTCTTCTACTACCTCTCG 375
            ||||||||||||| ||||| |||| || || | |||||| |||||| || || ||| ||
Db   244 ACATCATGCTTACCAACGTCCTTGACGCCTGACGCTGCAGCGCCGGTGCTCATTATCTGTTTG 303

Qy   376 GCTTCGCCTTCGCTTTCGGCTCCCCCTCCAAGGGCTTCATCGGTAAACATTTCTTCGGCC 435
            ||||||| ||||||||| || || | |||| |||||||||| ||||| || |||||||
Db   304 GCTTACGCCTTTGCGATCTCCGTCCAATGGTTCATCGGTAAACACTACTTTGGTC 363

Qy   436 TCAAGGACATCCCTTCATCCTCCTACGACTACAGCTACTTCCTACCAAATGGCCTTCG 495
            || |||| |||||| ||| ||| |||| ||||||||| ||| |||||||||||
Db   364 TCAAAGACATCCCACGGCCTTCTGCTGACTACTCCAACTTTCTCACCAAATGGCCTTTG 423

Qy   496 CCATCGCCGCCGCCGGCATCACCAGCGGAGGAAGCATCGCCGAACGACACAGTTCGTCGGCCT 555
            |||||||||||||||||| |||||||| || | |||| |||||||||||| ||||| ||||
Db   424 CAATCGCTGCGGCTGGAATCACAAGTGGCTCGATCGCTCGAACGGACACAGTTCGTCGGCTT 483

Qy   556 ATCTCATCTACTCCTCCTTCCTCACCGGCTTCGCTCATCGGTGGTCTCCCACTGGTTCT 615
            ||||| ||| | || || ||| || ||||| |||
Db   484 ACCTAAGTCATTCCTCTTTCTCTAACCGGGTTTGTTTTACCCGGTCTCTCACTGGTTCT 543

Qy   616 GGTCCCCAGACGGCTGGGCCTTCTGCCTT---TAAGATCACCGACCGGCTATTTCCACCG 672
            ||| | | ||| || || || || || | ||||| |||| |||||| |||||||||||
Db   544 GGTCAGTTGATGATGGGCCAGCCCGTTCCGTACCGATGGAGATTTGCTTTCAGCAGCCG 603

Qy   673 GCGTAATAGACTTCGCCCGTTCCGGCGTAGTCCACATGGTCGGCGGAATAGCCGGCCTAT 732
            || | | | |||||| ||| |||| | ||| || ||| |||| ||| | | |||||||
Db   604 GAGCGATAGATTTCGCTGGGTCCGTGTGTGTTCATATGGTCGGAGTATCGCTGGACTCT 663

Qy   733 GGGGAGCGGCTGATCGAAGGCCCAAGAATGGAGCGTTCGATCATGCAGGACGAGCGTGTGG 792
            ||||||| ||||| |||||| |||| ||||||||| |||| |||||||| || || || ||
Db   664 GGGGTGCGGCTCATCGAAGGTCGCCACGACTTGGCCGGTTCGATAACGGAGCGCCGTGCCATCG 723

Qy   793 CCTTGCGAGGCCACAGCGCGTCCGTTCCTAGTAGTCCTGGGAACCTTCTTGCTTTGGTTCGGTT 852
            || || | || || |||||| ||| |||||| | ||||||||| || ||| |||||||| ||
Db   724 CTCTTCGTGGCCACTCGGCGTCACTTGTTCCTTGGAACATTCCTCCTCTGGTTTGGAT 783

Qy   853 GGTACGGATTTAACCCCGGTTCATTAACAAAATCCTACTTACTTACGGTAACTCAGGAA 912
            ||||| | || ||||||||| | |||| | |||| || |||||| |||| |||||||
Db   784 GGTACGGATTTAACCCCGGTTCCTTCAACAAGATCCTAGTCACGTACGAGA--CAGGCA 840

Qy   913 ATTACTACGGTCAATGGAGCGCCGGTTGGCAGAACCGCGCGTCACCACACCTAGCGGGGT 972
            | || || ||||||||||||||||| |||| |||| ||||| ||||| || || ||| |||

Db   841 CATACAACGGCCAGTGGAGCGCCGGTCGGCGACGGACAGCTGTCACAACACGTTAGCTGGCT 900

Qy   973 CAACAGCTGCTTGACCACCGCTATTCGTAACGGGTAAACGGGTGATATCCGGTCACTGAACGTGA 1032
            || ||||||| ||||| | | || ||| ||||||||| |||||||| |||| |||||| |||
Db   901 GCACCGGCGCGCCGACACAACCCTATTTGGGAAACGTCACTCCTCGGAGACATTGGAACGTCA 960

Qy   1033 CCGATGCTGCAACGGGCTGTTAGGCGGTTCGCGGCGGCCGATAACAGCCGGTTGCTTCCGTGG 1092
            || || || | ||| ||| ||| || ||||||||| |||| ||| ||||| ||||| ||||
Db   961 CTGATGATGCAACGGCCTCCGGAGGGTTGCAGCCATAACTGGTGGCTCGCTCGTCG 1020

Qy   1093 TTGAGCCATGGGCAGCCATCGTATGCGGTTTTGTTGCTTCCTATAGTATTAATAGCTTGCA 1152
            ||||||||||||| |||| ||| ||||| ||| | ||| || | | |||||| |||||
Db   1021 TTGAGCCATGGGCTGCGATCATCTGCGGGTTCGTGCGTGGCGGCCCTAGTCCTCCGGATGCA 1080

Qy   1153 ACAAATTAGCAGAGAGAGGTTAAGTTGCACGATCCTCGGAGGCGCGGCAGTTGCACGGTG 1212
            |||| | | || | ||| ||||| |||||| ||||| ||| || ||||| | || ||
Db   1081 ACAGCTCCGCTGAGAGCTCAAATACGACGACCCCTTCTTGAGGCAGCAACTACACGGTG 1140

Qy   1213 GGTGTGGCACGTGGGGGGTGATATTCACGGCGGTGTGTTCGCAAAAAAGGAGTATGTGAAGG 1272
            ||||||| ||||| || |||| |||| ||||| |||| | ||||| || |||| |||||
Db   1141 GTTGCGGTGCGTGGGGGGACTAATATTCACGGGCTCTCGCTCGCTGCAAGTACTTGAACC 1200

Qy   1273 AGGTTTACGGGTTG------GGGAGGGCGCACGGGTTGCTCATGGGGGTGGTGGGAAGT 1326
            | | | |||||||| |||| | || ||| || |||| ||| |||| |||||
Db   1201 AGATTTACGGCAACAAACCCGGAGGCCACACGGTTTGTTTTATGGGGTGGAGTGGGAAAAC 1260

Qy   1327 TGCTGCGGCGGCACGTGATTCAGATTCTGGTGATTGGTCGGTGGGTGAGTGCGGACCATGG 1386
            |||| |||| |||| || | ||| || |||| ||||| | |||| || |||| ||||||
Db   1261 TACTTGGAGCTCAGCTGATTCAGATCATTGTGATCACGGGTTTGGGTAAGTGCGACCATGG 1320

Qy   1387 GACCCTTGTTTTGGGGGTTGAATAAACTGAAGCTGTTGAGGATTCTTCAGAGGATGAGC 1446
            |||| | |||| || || || ||| || ||||| |||| ||||| |||||||||||
Db   1321 GGACACTTTTCTCATCCTCAAGAAAATGAAAATTGTTGCGGATATCGTCGGAGGATGAGA 1380

Qy   1447 TTGCGGGATGGACATGACTCGCCATGGAGGCTTTGCTTATGGCTTATGAGGATGATGAGA 1506
            |||| |||| |||| | || ||||| || || |||| ||| | ||| | ||||||||
Db   1381 TGGCCGGTATGGATATGACCAGCAGCACGTGGTGTTTCCTTATATGTACTTTGATGATGTG 1440

Qy   1507 CGCACAAGCACATGGACATGACTCGCCATGGAGGAGGGTTGGGCCCACCACCACCA 1566
            ||| ||||| ||||| ||||| || ||||| | || | ||| |||| |||||||
Db   1441 AGTCTCACAAAGCCATTCAGCTTAGGAGAGTTGAGCCACGGATTTGAGCCACGACCATTCTTCTTCTTCTGTGTG 1500

Qy   1567 CTGAT 1571
            || ||
Db   1501 CTAAT 1505

RESULT 8
AR399484
LOCUS   AR399484   1748 bp   mRNA   linear   PAT 18-DEC-2003
DEFINITION Sequence 1 from patent US 6620610.
ACCESSION   AR399484
VERSION   AR399484.1 GI:40141532
KEYWORDS   .
SOURCE   Unknown.
  ORGANISM Unknown.
            Unclassified.
REFERENCE   1 (bases 1 to 1748)
  AUTHORS   Frommer,W.-B. and Ninnemann,O.
  TITLE   DNA sequence from Arabidopeis thaliana encoding ammonium
            transporter, and plasmids, bacteria and yeast comprising the DNA
            sequence
  JOURNAL   Patent: US 6620610-A 1 16-SEP-2003;
FEATURES   Location/Qualifiers
     source   1..1748
                     /organism="unknown"
                     /mol_type="mRNA"
ORIGIN

Query Match   44.3%;   Score 755.8; DB 6;   Length 1748;
Best Local Similarity 71.7%; Pred. No. 5.4e-124;
Matches 1036; Conservative   0; Mismatches 397; Indels 12; Gaps   3;

```
Qy   1213 GGTGTGGCACGTGGGGGTGATATTCACGGCGTGTTGCGCAAAAAGGAGTATGTGAAGG 1272
          ||||||||||||||||||| ||||||||||||||||||||  ||||||||||
Db   1141 GTTGCGGGTGCGTGGGGACTAATATTCACGGCTCTCTTCGCTCAGAAAAGTACTTGAACC 1200

Qy   1273 AGGTTTACGGGTTG------GGGAGGGGCGCACGGGTTGCTCATGGGGGGTGGTGGGAAGT 1326
          |||||||||||||      |||||||||||||||||||||||||||||||||||||||
Db   1201 AGATTTACGGCAACAAACCCGGAGGGCCACACGGTTTGTTTATGGGCGGTGGAGGAAAAAC 1260

Qy   1327 TGCTGGCGGCCACGTGATTCAGATTCTGTGATTGCTGGGTGGTGGTAGTGCGACCATGG 1386
          ||||||||||||||||||||||||||||||||||| |||||| ||||||||||||
Db   1261 TACTTGGAGTCAGCTGATTCAGATCATTGTGATCACCGGGTTGGGTAAGTGCGACCATGG 1320

Qy   1387 GACCCTTGTTTTGGGGTTGAATAAACTGAAGCTGTTGAGGATTTCTTCAGAGGATGAGC 1446
          ||||||||||||||||||||||||| |||||||||||| |||||||||||||||
Db   1321 GGACACTTTCTTCATCCTCAAGAAAATGAAATTGTTGCGGGATATCGTCCGAGGATGAGA 1380

Qy   1447 TTGCGGGGATGGACATGACTCGCCATGGAGGCTTTGCTTATGCTTATGAGGATGATGAGA 1506
          |||||||||||||||||||||||||||| ||||||||||||||| |||||||||
Db   1381 TGGCCGTATGGATATGACCAGGCACGGTGGTTTGCTTATATGTACTTTGATGATGATG 1440

Qy   1507 CGCACAAGCATGGATGCAGTTGAGGAGGGTTGGGCCCAACGGCGTCTTCCACACCCACCA 1566
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 AGTCTCACAAAGCCATTCAGCTTAGGAGAGTTGAGCCACGACGATCCTCCTTCTCTGGTG 1500

Qy   1567 CTGAT 1571
          || ||
Db   1501 CTAAT 1505
```

RESULT 9
```
ATT6G15
LOCUS       ATT6G15             113970 bp    DNA     linear   PLN 20-APR-1999
DEFINITION  Arabidopsis thaliana DNA chromosome 4, BAC clone T6G15 (ESSA
            project).
ACCESSION   AL049656
VERSION     AL049656.1  GI:4678371
KEYWORDS    .
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1
  AUTHORS   Bevan,M., Murphy,G., Ridley,P., Hudson,S., Bancroft,I., Mewes,H.W.,
            Mayer,K.F.X. and Schuelle;.C.
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 113970)
  AUTHORS   EU Arabidopsis sequencing.project.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-APR-1999) MIPS, at the Max-Planck-Institut fuer
            Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail:
            schuelle@mips.biochem.mpg.de, mayer@mips.biochem.mpg.de Project
            Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge
            Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK,
            E-mail: michael.bevan@bbsrc.ac.uk
COMMENT     Information on performance of analysis and a more detailed
            annotation of this entry and other sequences of chromosomes 3, 4
            and 5 can be viewed at: http://www.mips.biochem.mpg.de/proj/thal/.
FEATURES             Location/Qualifiers
     source          1..113970
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /variety="Columbia"
                     /db_xref="taxon:3702"
                     /chromosome="4"
     misc_feature    1..9204
                     /note="overlap to BAC T9E8; please refer EMBL acc:AL049608
                     for analysis and annotation"
     gene            9187..11139
                     /gene="T6G15.10"
     gene            complement(9187..11139)
                     /gene="T6G15.10"
     CDS             complement(9187..11139)
```

```
Qy   136  ACGGCTTCGCCGCCGGCCTCCCTTATCTGCGGCCATTTCGCCGCGTGACAGGAAGTTCG 195
          |||||||||||||||||||||||||||| ||||||||||||||||||||||
Db   64   ATGCCACGGCGGCGGCCAACTACACATGTGGCCAGCTAGGCGACGTCAACAACAAATTCA 123

Qy   196  TCGACACGGCCTTCGCGTCGACACACCTACCTCCTCCTTTCGCCTACCTCGTTTTTT 255
          |||||||||||||||| ||||||||||||||||| |||||||||||||
Db   124  TCGACACCGCTTTGCCTTTGGATCTGCTATAGACAACACTTACCTCCTCCGCCTACCTTGTCTTCT 183

Qy   256  CTATGCAGCTCGGCTTCGCCATGCTCTGCGCCGGCCTCCGGCCGGCCAAGAACACCATGA 315
          |||||||||||||||||||||||| ||||||||||||||||||||||
Db   184  CTATGCAGCTTGGCTTCGCTATGCTCGTGTGCGCGGTTCGCGGTGGAGCCAAGAACATCATGA 243

Qy   316  ACATCATGCTCACCAACGTCCTGGACGCGTGCCGCCGGCCGCCTCTTCTACTACCTCTTCG 375
          |||||||||||| ||||||||||||||||||||||||||||||||
Db   244  ACATCATGCTTACCAACGTCCTTGACGCTGCAGCCGGTCGTTCTCTATTATCGTTTG 303

Qy   376  GCTTCGCCCTTCGCTTCGGCTCCCCTCCAAGCGCTTCATCGTAAACATTTCTTCGGCC 435
          ||||||||||| |||||||||||||||||||||||||||| |||||||||
Db   304  GCTACGCCTTTGCCTTTGGATCTCGGTCCAATGGTTTCATCGTAAACACTACTTTGGTC 363

Qy   436  TCAAGGACATCCCTTCATCCTCCTACGACTACAGCTACTTCCTCTACCAATGGCCTTCG 495
          |||||||||| |||||||||||||||||||||||| |||||||||||||||
Db   364  TCAAAGACATCCCCACGGCCTCTGCGTTGGACTACTCCAACTTTCTCTACCAATGGCCCTTTG 423

Qy   496  CCATCGCCGCGCCGGCATCACCAGCCGGGAAGCATCGCCGAACGCACACAGTTCGTCGGCCT 555
          |||||||||||||||||||||||||||||||| ||||| |||||
Db   424  CAATCGCTGCGGCTGGAATCACAAGTGGCTCGATCGCTGAACGGACACAGTTCGTCGGCTT 483

Qy   556  ATCTCATCTACTCCTCCTTCCTCACCGGCTTCGTCTATCCGGTGCGTGTGCTGTCCCACTGTTTCT 615
          |||||||||||||| |||| ||||||||||| ||||||||| |||||||
Db   484  ACCTAATCTATTCCTCTCTTTCTTAACCGGGTTTGTTTACCCGGTCGTCTCTCACTGGTTCT 543

Qy   616  GGTCCCCAGACGCGCTGGGCCTTCTGCCTT---TAAGATCACCGACCGGCTATTTCCACCG 672
          ||||||||||||||||||||||| ||||| ||| ||||| |||
Db   544  GGTCAGTTGATGGGCCGCAGCCCCGTTCCGTACCGATGGAGATTTGCTTTTCAGCACCG 603

Qy   673  GCGTAATAGACTTCGCCGGTTCCGGCGTAGTCCACATGGTCGGCGGAATAGCCGCCTAT 732
          |||| ||||||| |||||||||||||||||||||||||||||||||||
Db   604  GAGCCGATAGATTTCGCTGGGTCCCGGTGTGTTCGGAGGTATCGGAGGTATCGCTGGACTCT 663

Qy   733  GGGGAGGCGTGATCGAAGGCCCAGAAGAATGGACGTTTCGATCATCGAGACGAGAGCTGTGG 792
          ||||||||||||||| |||||||||||||||||||||||||||||
Db   664  GGGGTGCGCTCATCGAAGGTCACGACTTGGCCGGTTCGATAACGGAGGCCGTGCCATCG 723

Qy   793  CCTTGCGAGGCCACAGCCGCGTCCTTAGTAGTCCTGGGAACCTTCTGCTTTGGTTCGGTT 852
          |||||||||||||||||||| ||||| |||||||||||||||||||||||
Db   724  CTCTTCGTGGCCACTCGCGCGTCACTTGTTGTCCTTGGAACATTCCTCCTCGGTTTGGAT 783

Qy   853  GGTACGGATTTAACCCCGGTTCATTTAACAAAATCCTACTTACGGTAACTCAGGAA 912
          |||||||||||||||||||| ||||| |||||||||||||||| |||||
Db   784  GGTACGGATTTAACCCCGGTTCCTTCACAAGATCCTACCTCACGTACGAGA---CAGGCA 840

Qy   913  ATTACTACGGTCAATGAGAGCGCGGTTGGCAGAACCGCGGTCACCACTACCCTAGCGGGGT 972
          ||||||||||||| |||||||||||||||||||| ||||||||||||
Db   841  CATACAACGGCCAGTGAGCAGCCAGTCGGACGACGACAGCGTCACACAACGTTAGCTGGGCT 900

Qy   973  CAACAGCTGCCTTGACCACGCTATTCGGTAAACGGGTGATATCCGGTCACTGGAACGTGA 1032
          |||||||||| |||||| ||||||||| |||||||||||||||
Db   901  GCACCCGCGGCCTGACAACCCTATTTGGAAACGTCTACTCTCGGGACATTGGAACGTCA 960

Qy   1033 CCGATGTCTGCAACGGCGTGTTGAGGCCGGTTCGCGGGATAACAGCCGGTTGCTCCGTGG 1092
          |||| ||||||||||||||| ||||||||||||||||||||||||
Db   961  CTGATGTATGCAACGGCCTCCTCGGAGAGGTTTGCAGCCATAACTGGTGCTGCTCGTGTCG 1020

Qy   1093 TTGAGGCCATGGGCAGCCATCGTATAGCCGGTTTTGTTGTTGCTTCTATAGTAGTTAAATAGCTTGCA 1152
          ||||| |||||||||||||||||||||||||||| |||||||||||
Db   1021 TTGAGCCATGGGCTCGGATCATCTCGCGGGTTCGTCGGCGGCCCTAGTCCTTCCTCCGGATGCA 1080

Qy   1153 ACAAATTAGCAGAGAGAGGTTAAGTTCGACGATCCTCTGGAGGCGGCGCAGTTGCACGGTG 1212
          |||| ||| ||||||||||||| ||||||||||||||||||||||||||
Db   1081 ACAAGCTCGCTGAGAGAGCTCAAATACGACGAGCCCTCTTGAGGCAGCAACTACACGGGTG 1140
```

```
                     /gene="T6G15.10"
                     /note="similarity to G9A protein, Mus musculus, AF109906"
                     /codon_start=1
                     /product="putative protein"
                     /protein_id="CAB41104.1"
                     /db_xref="GI:4678372"
                     /db_xref="SWISS-PROT:Q9TOG7"
                     /translation="MGSSHIPLDPSLNPSPSLIPKLEVTESTQNLAFQLPNTNPQAL
           ISSAVSDFNEATDFSSDYNTVAESARSAFAQRLQRHDDVAVLDSLTGAIVPVEENPEP
           EPNYSTSDSSPSVATQRPRPQPRSSELVRITDVGPESERQFRRHVRKTRMIYDSLRM
           FLMEEAKRNGVGGRARADGKAGKAGSMMRDCMLWMNRDKRIVGSIPGVQVGDIFFF
           RFELCVMGLHGHPQSGIDFLTGSLSSNGEPIATSVIVSGGYEDDDDQDVIMYTGQGG
           QDRLGRQAEHQRLEGGNLAMERSMYYGIEVRVIRGLKYENEVSSRVYVYDGLFRIVDS
           WFDVGKSGFGVFKYRLERIEGQAEMGSSVLKFARTLKTNPLSVRPRGYINFDISNGKE
           NVPVYLFNDIDSDQEPLYYEYLAQTSFPPGLFVQQSGNASGCDCVNGCGSGCLCBAKN
           SGEIAYDYNGTLIRQKPLIHECGSACQCPPSCRNRVTQKGLRNRLEVFRSLETGWGVR
           SLDVLHAGAFICEYAGVALTREQANILTMNGDTLVYPARFSSARWEDWGDLSQVLADF
           ERPSYPDIPPVDFAMDVSKMRNVACYISHSTDPNVIVQFVLHDHNSLMFPRVMLFAAE
           NIPPMTELSLDVGVVDDWNAKLAICN"
exon       complement(9187..11139)
                     /gene="T6G15.10"
                     /number=1
gene       12507..14016
                     /gene="T6G15.20"
gene       complement(join(12507..12811,13043..13327,13836..14016))
                     /gene="T6G15.20"
CDS        complement(join(12507..12811,13043..13327,13836..14016))
                     /gene="T6G15.20"
                     /note="similarity to various predicted proteins,
                     Arabidopsis thaliana"
                     /codon_start=1
                     /product="putative protein"
                     /protein_id="CAB41105.1"
                     /db_xref="GI:4678373"
                     /db_xref="SPTREMBL:Q9TOG8"
                     /translation="MMMEVLVARGGCDHGDGNDVVVMVVAVVVVTARWWLWLWL
           RLASYDGCDGDGGCGGTSFWYDNWCSLGSLYEVLGARGCIDMGISSSWSVADVLRSH
           VRRHRTATLNSIEDEIDALRQRHDPTCDIPLWKCCNGSFKHKFFSRDTWYEIPLTRD
           MLGSRYTTRWDSLVDLSTSSALTQIHMFVFRYVFQLTTHSLWKERNGRRHGSSPIPAP
           SLAKMIDKNMRNRLSTLATSGSPIYKSGLRYWFSTHCR"
exon       complement(12507..12811)
                     /gene="T6G15.20"
                     /number=1
intron     complement(12812..13042)
                     /gene="T6G15.20"
                     /number=1
exon       complement(13043..13327)
                     /gene="T6G15.20"
                     /number=2
intron     complement(13328..13835)
                     /gene="T6G15.20"
                     /number=2
exon       complement(13836..14016)
                     /gene="T6G15.20"
                     /number=3
gene       21209..22187
                     /gene="T6G15.30"
CDS        join(21209..21323,21417..21546,21647..22187)
                     /gene="T6G15.30"
                     /note="similarity to snapdragon myb protein 305 homolog,
                     Arabidopsis thaliana, U62743
                     Contains Myb DNA-binding domain repeat signatures, Myb_1
                     [WTABEDRLL], ATP/GTP-binding site motif A (P-loop)
                     [GLKRNGKS]"
                     /codon_start=1
                     /product="myb-like protein"
                     /protein_id="CAB41106.1"
                     /db_xref="GI:4678374"
                     /db_xref="GOA:Q9TOG9"
                     /db_xref="SPTREMBL:Q9TOG9"
                     /translation="MVEEVWRKGPWTAEBDRLLIEYVRVHGEGRWNSVSKLAGLKRNG
           KSCRLRWVNYLRPDLKRGQITPHEESIILELHAKWGNRWSTIARSLPGRTDNEIKNYW

           RTHFKKAKPTTNAEKIKSRLLKRQHFKEQREIELQQEQLFQFDQLGMKKIISLLE
           ENNSSSSDGGDVFYYPDQITHSSKPFGVNSNSLEEQLQGRFSPVNIPDANTMNEDN
           AIWDGFWNMDVVNGHGGNLGVVAATAACGPRKPYFHNLVIPFC"
exon       21209..21323
                     /gene="T6G15.30"
                     /number=1
intron     21324..21416
                     /gene="T6G15.30"
                     /number=1
exon       21417..21546
                     /gene="T6G15.30"
                     /number=2
intron     21547..21646
                     /gene="T6G15.30"
                     /number=2
exon       21647..22187
                     /gene="T6G15.30"
                     /number=3
gene       23671..24818
                     /gene="T6G15.40"
gene       complement(join(23671..24069,24179..24312,24671..24818))
                     /gene="T6G15.40"
CDS        complement(join(23671..24069,24179..24312,24671..24818))
                     /gene="T6G15.40"
                     /note="similarity to predicted protein, Arabidopsis
                     thaliana"
                     /codon_start=1
                     /product="putative protein"
                     /protein_id="CAB41107.1"
                     /db_xref="GI:4678375"
                     /db_xref="SPTREMBL:Q9TOH0"
                     /translation="MSLESDIKESKRVVEPFDDDCAVCLEPLANDADERTVVKLRCS
           HKFHLDCVGSSFNIKNKMECPCCRQIEKGKWLFAEPVDQLEEDDMSPLDMLTNMVSPS
           FLHIDDGIFTNIASVLFPNALNQISNNANVSSNPLGEPRNTLLDTRMQRQVDEHVAST
           VDSERRLPDMLFELHDRTGGSGPSGYGRERSGGEGGENSCGGEAVABEDQVGEAAEGA
           EMMDNLNF"
exon       complement(23671..24069)
                     /gene="T6G15.40"
                     /number=1
intron     complement(24070..24178)
                     /gene="T6G15.40"
                     /number=1
exon       complement(24179..24312)
                     /gene="T6G15.40"
                     /number=2
intron     complement(24313..24670)
                     /gene="T6G15.40"
                     /number=2
exon       complement(24671..24818)
                     /gene="T6G15.40"
                     /number=3
misc_feature   27751..28154
                     /note="ESTs T76021 and AA712162 match to positions
                     27751-27983 and 28079-28154; no obvious genmodel"
misc_feature   30974..31279
                     /note="EST AA404812 matches to position 30974-31279; no
                     obvious genmodel"
gene       41521..42408
                     /gene="T6G15.50"
gene       complement(join(41521..41665,42112..42181,42246..42408))
                     /gene="T6G15.50"
CDS        complement(join(41521..41665,42112..42181,42246..42408))
                     /gene="T6G15.50"
                     /note="contains EST gb:AA605343"
                     /codon_start=1
                     /product="putative protein"
                     /protein_id="CAB41108.1"
                     /db_xref="GI:4678376"
                     /db_xref="SPTREMBL:Q9TOH1"
                     /translation="MAAKFRISSSSFSHRASDSTSSSSSYSSLLALPQFFCPPSPLG
           FPEFKLHAKLGGGDGEVKPKDKKKFITKEEBEPQYWQSVGERSGGENPMKTPLPYIIIF
           GMSTPFVILAIAPANGWIKVPTR"
exon       complement(41521..41665)
```

```
            /gene="T6G15.50"
            /number=1
intron      complement(41666..42111)
            /gene="T6G15.50"
            /number=1
exon        complement(42112..42181)
            /gene="T6G15.50"
            /number=2
intron      complement(42182..42245)
            /gene="T6G15.50"
            /number=2
exon        complement(42246..42408)

Query Match  44.3%;  Score 755.8;  DB 8;  Length 113970;
Best Local Similarity  71.7%;  Pred. No. 5.5e-124;
Matches 1036;  Conservative  0;  Mismatches 397;  Indels  12;  Gaps  3;

Qy    136 ACGCCTCCGCGCCGCCTCCCTATCTGCGCCCATTCGCGCCGTGGACAGCAAGTTCG 195
          |||||  ||  | |||  |||  ||||  || ||||  | | |||||||
Db  42797 ATGCCACGGCGGCGGCCGGCCAACTACATCTGTGGCCAGTTAGGGACGTCAACAACAAACAAATTA 42856

Qy    196 TCGACACGGCCTTCGCGTGCGACAACACCTACCTCCTCTTTCGCCTACCTCGTTTTTT 255
          |||  |||| | |||| |||||   |||   || |||| ||| || ||||| || |
Db  42857 TCGACACCCCTTTGCTATAGACACACTTACCTTCTCTTTCTCCTCCTTGTCTTCT 42916

Qy    256 CTATGCAGCTCGGCTTCGCCATGCTCGCGCCGGCTCCGTCCGCGCCAAGAACACCATGA 315
          |||||  ||||| ||| | ||| ||  |||||  ||  || |||||||||| ||
Db  42917 CTATGCAGCTTGGCTTCGCTATGCTCGTGTGCCGGGTTCCGTGAGAGCCAAGAATACTATGA 42976

Qy    316 ACATCATGCTCACCAAAGTCCTGGACGCTGCCGCGGCGGGCTCTTCTACTACCTCTTCG 375
          ||||||||| || | |||  ||||   || |||||||||||||| | ||  |  || ||
Db  42977 ACATCATGCTTACCAAACGTCCTTGACGCTGCAGCGCGTGCTGGTGTCTCTTCTATTATCTGTTTG 43036

Qy    376 GCTTCGCCTTCGCTTCGCCTTCGGCTCCCCCTCCAAACGGCTTCATCGGTAAACATTTCTTCCGCC 435
          ||| |||| ||||| || |||| || | | | ||||||| ||||| ||| |||  ||||| ||| |
Db  43037 GCTACGGCCTTTGCCTTTGGATCTCCGTCCAATGGTTTCATCGGTAAAACACTACTTTGGTC 43096

Qy    436 TCAAGGACAATCCCTTCATCCTCCTACGACTACAGCTACTTCCTCTACCAATGGGCTTCG 495
          |||| ||| |||||  || ||||||| | | |   || ||||| ||| ||| |||||| |
Db  43097 TCAAAGACATCCCACGGCCTCTGCTGACTACTCCAACTTTCTCTACCAAATGGGCCTTTG 43156

Qy    496 CCATCGCGCCGCGGCCGCATCACCAGCGGAAGCATCGCCGAACGCACACAGTTCGTGGCCT 555
          ||||| | | || | |  |||  |  ||| || ||||| ||||| |||||||| |||| ||
Db  43157 CAATCGTGCCGCGGTGGTGGAATCACAAGTGGCTCGATCGCTGAACGGACACAGTTCGTTGGCTT 43216.

Qy    556 ATCTCATCTACTCCTCCTTCCTCACCGGCTTCGTCTATCCGGGTGGTCTCCCACTGTTCT 615
          || || ||||  |||||| || |||| |||| ||| || ||||  ||| |||||||||||
Db  43217 ACCTAATCTATTCCTCTTTTCTTAACCGGTTTGTTTACCCGGTCGTGTCTCTCACTGTTCT 43276

Qy    616 GGTCCCCAGACGGCTCTGGCGCCTTCGCCCTT---TAAGATCACCGACCGGCTATTTTCCACCG 672
          |||  |  | | | ||| || || ||| | ||||   ||||  ||| || |||||| |||||
Db  43277 GGTCAGTTGATGGAGTGGGCCAGCCCGTTCCGCTACCGATGGAGATTTGCTTTTCAGCACCG 43336

Qy    673 GCGTAATAGACTTCGCCCGGTTCCGGGTAGTCCCGGTAGTCGGCGGAATAGCCGGCCTAT 732
          || | || ||| | ||||  ||  |   | || | || ||  | || ||||| ||| | |
Db  43337 GAGCGATAGATTTCGGTGGTCCGGTGTTGTTCATATGGTCCGGAGGTATCGCTGCTGGACTCT 43396

Qy    733 GGGGAGGCGCTGATCGAAAGCCAAGAATGGACGGTTCGATCATCGAGGACGAGCTGTGG 792
          ||||| ||| ||  | || |  |||||||| ||||| || ||||| | || |||| |||
Db  43397 GGGGTGCGCTCATCGAAGGTCCACGACGTTGGGCCGGTTCGATAACCGGAGGCGGTGCCATCG 43456

Qy    793 CCTTGCGAGGCCACAGCGCGTCCTTAGTAGTCCTCGGGAACCTTCTTGCTTTGGTTCGGTT 852
          || |  || ||| ||||| ||  |||||||  || ||||| ||||  |  |||| ||| ||
Db  43457 CTCTTCCGTGCGCCACTCGCGTCACTTGTTGTCCTTGGAACATTCCTCCTCGTTGTTGGAT 43516

Qy    853 GGTACGGCGATTTAACCCCGGTTCATTTAACAAAATCCTACTTACTTACGGTAACTCAGGAA 912
          |||| |||||| | ||||||||  ||| | || | | ||  ||| || |
Db  43517 GGTACGGCGATTTAACCCCGGTTCCTTCAACAAGATCCTAGTCACGTACGAGA---CAGGCA 43573

Qy    913 ATTACTACGGTCAATGGAGAGCGCGGTTGGACGAACCGGCGTCACCACTACCCTAGCGGGGT 972
          || |  ||||| ||  || | || ||  | |  | ||||||  | |||  | |||| || ||
Db  43574 CATCACAACGGCCAGTGAGCGCGGTCGGACGACGAGCAGCTGTCACAACAACAAGTTAGCTGGCT 43633
```

```
Qy    973 CAACAGCTGCCTTGACCACGCTATTCGGTAAACGGGTGATATCCGGTCACTGGAACGTGA 1032
          |||||  ||  ||||   ||||||||||||  |||| ||| | |||| | ||||| | |
Db  43634 GCACCGCGGCGGCTGACAACCCTATTTGGGAAACGTCTACTCTCCGGGACATTGGAACGTCA 43693

Qy   1033 CCGATGTCTGCAACGGCGTGTTAGGCGGTTCGGCGGCATAACAGCCGGTTGCTCCGTGG 1092
          ||| | |  |||| ||| ||||  ||||||  || || || | ||||| |||||||| |
Db  43694 CTGATGTATGCAACGGCGTTTCGCAGCCCATAACTGGTGCGTCGTCTTGTCG 43753

Qy   1093 TTGAGGCCATGGCAGCCATCGTATGCGGGTGTTTGTTGCTTCTATAGTATTAATAGCTTGCA 1152
          ||||| |||||| |||| | | | ||| |||||  ||| ||||  | |  | | ||  || |
Db  43754 TTGAGCCATGGGCTGGCATCATCTGCGGGTCGTGCGGGCCCAGTCCTTCCTCCGATGCA 43813

Qy   1153 ACAAATTAGCAGAGAAGGTTAAGTTCGACGATCCTCTGGAGGCGGCGCAGTTGCACGGTG 1212
          ||| | ||| |  |  || |||  | ||   | |||| |||||  | ||  | || ||||
Db  43814 ACAAGTCGCTGAGAGCTCAAATACGACGACCCTTTCAGGCAGCAACAACTACACGGTG 43873

Qy   1213 GGTGTGGCACGTGGGGGGTGATATTCACGGCGTTGTTCGCAAAAAGAGTATGTGAAGG 1272
          |||  ||| |  ||| ||| ||||  |||  ||  ||  |  || |  | | | ||||
Db  43874 GTTGCGGTGCGTGGGGGGACTAATATTCACGGCTTCTCTTTCGCTCAAGAAAAGTACTTGAACC 43933

Qy   1273 AGGTTTACGGGTTG------GGGAGGGCCACGGGTTGCTCATGGGGGTGGTGGGAAGT 1326
          | ||||| || ||       ||| || || |||| || ||||  | || |||| |||
Db  43934 AGATTTACGGCAACACAAACCCGGAAGGCCACACGGTTGTTTATGGGCCGTGGAGGAAAAC 43993

Qy   1327 TGCTGGCGGCGCACGTGATTCAGATTCGGTGATTGCTGGGGTCGGTTAGTGCGACCATGG 1386
          ||| ||||| ||||| | | || || || |  |||  ||  |||| ||| |||| ||||
Db  43994 TACTTGGAGCTCAGCTGATTCAGATCATTGTGATCACGGGTTGGGTAAGTGCGACCATGG 44053

Qy   1387 GACCCTGTTTTGGGGGTTGAATAAACTGAAGCTGTTGAGGATTTCTTCAGAGGATGAGC 1446
          ||||  | | | | ||| | |  | |||  |  | | |  || ||| || |||||| || |
Db  44054 GGACACTTTCTTCATCCTCAAGAAAATGAAATTGTTGCGGATATCGTCCGAGATGATGAGA 44113

Qy   1447 TTGCGGGGATGGACATGACTCGCCATGGAGGCTTTGCTTATGCTTATGGATGATGAGA 1506
          |||| || ||| | |||  | | ||||||||  | |  | ||  || || |||| |||
Db  44114 TGGCCGGTATGGATATGACCAGGCACGGTGGTTTTGCTTTATATGTACTTTGATGATG 44173

Qy   1507 CGCACAAGCATGGGATGCAGTTGAGGAGGTTGGGCCCAACGCGTCTTCCACCACCACCA 1566
          ||||||| | |||||| | ||  ||  | ||  | ||||| | ||||||||  |  |||
Db  44174 AGTCTCACAAAGCCATTCAGCTTAGCAGAGTTGAGCCAGCACGCGATCTCCTTCTCCTTCTGTGG 44233

Qy   1567 CTGAT 1571
          ||  |
Db  44234 CTAAT 44238
```

```
RESULT 10
ATCHRIV36
LOCUS       ATCHRIV36        199634 bp    DNA       linear    PLN 16-MAR-2000
DEFINITION  Arabidopsis thaliana DNA chromosome 4, contig fragment No. 36.
ACCESSION   AL161536
VERSION     AL161536.2  GI:7268028
KEYWORDS    .
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 46786; 47571 to 111411)
  AUTHORS   Pohl,T., Weizenegger,T., Mewes,H.W., Lemcke,K. and Mayer,K.F.X.
  JOURNAL   Unpublished
REFERENCE   2  (bases 46394 to 47589)
  AUTHORS   Robben,J., Grymonprez,B., Volckaert,G., Mewes,H.W., Lemcke,K. and
            Mayer,K.F.X.
  JOURNAL   Unpublished
REFERENCE   3  (bases 102208 to 102209)
  AUTHORS   Murphy,G., Ridley,P., Hudson,S., Mewes,H.W., Lemcke,K. and
            Mayer,K.F.X.
  JOURNAL   Unpublished
REFERENCE   4  (bases 1 to 199634)
  AUTHORS   EU Arabidopsis sequencing project.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-MAR-2000) MIPS, at the Max-Planck-Institut fuer
            Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail:
```

```
COMMENT     lemcke@mips.biochem.mpg.de,mayer@mips.biochem.mpg.de Project
            Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge
            Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK,
            E-mail: michael.bevan@bbsrc.ac.uk
            Information on performance of analysis and a more detailed
            annotation of this entry and other sequences of chromosomes 3, 4
            and 5 can be viewed at: http://www.mips.biochem.mpg.de/proj/thal/
            this fragment has an overlap with ATCHRIV35 at the 5' end and an
            overlap with ATCHRIV37 at the 3' end.
FEATURES             Location/Qualifiers
     source          1..199634
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /variety="Columbia"
                     /db_xref="taxon:3702"
                     /chromosome="4"
     repeat_region   7835..7865
     gene            8581..10357
                     /gene="AT4g13260"
     gene            complement(join(8581..8778,8909..9031,9134..9382,
                     9680..10357))
                     /gene="AT4g13260"
     CDS             complement(join(8581..8778,8909..9031,9134..9382,
                     9680..10357))
                     /gene="AT4g13260"
                     /note="similarity to dimethylaniline monooxygenase
                     (N-oxide-forming) (EC1.14.13.8) -Oryctolagus cuniculus,
                     PID:g164989
                     Contains Prokaryotic membrane lipoprotein lipid attachment
                     site AA33-43"
                     /codon_start=1
                     /product="putative protein"
                     /protein_id="CAB78368.1"
                     /db_xref="GI:7268029"
                     /db_xref="GOA:Q9SVQ1"
                     /db_xref="SPTREMBL:Q9SVQ1"
                     /translation="MEFVTETLGKRIHDPYVBETRCLMIPGPIIVGSGPSGLATAACL
                     KSRDIPSLILERSTCIASLWQHKTYDRLRLHLPKDFCELPLMPFPSSYPTYPTKQQFV
                     QYLESYABHFDLKPVFNQTVBEAKFDRRCGLWRVRTTGGKDETMEYVSRWLVVATGE
                     NAEEVMPEIDGIPDFGGPILHTSSYKSGEIFSEKKILVVGCGNSGMEVCLDLCNFNAL
                     PSLVVRDSVHVLPQEMLGISTFGISTSLLKWFPVHVVDRFLLRMSRIVLGDTDRLGLV
                     RPKLGPLERKIKCGKTPVLDVGTLAKIRSGHIKVYPELKRVMHYSABFVDGRVDNFDA
                     IILATGYKSNVPMWLKGVNMFSEKDGFPHKPFPNGWKGESGLYAVGFTKLGLLGAAID
                     AKKIAEDIEVQRHFLPLARPQHC"
     exon            complement(8581..8778)
                     /gene="AT4g13260"
                     /number=1
     intron          complement(8779..8908)
                     /gene="AT4g13260"
                     /number=1
     exon            complement(8909..9031)
                     /gene="AT4g13260"
                     /number=2
     intron          complement(9032..9133)
                     /gene="AT4g13260"
                     /number=2
     exon            complement(9134..9382)
                     /gene="AT4g13260"
                     /number=3
     intron          complement(9383..9679)
                     /gene="AT4g13260"
                     /number=3
     exon            complement(9680..10357)
                     /gene="AT4g13260"
                     /number=4
     tRNA            17702..17782
                     /product="tRNA-Leu"
                     /note="codon recognized: CUU"
     gene            18021..18599
                     /gene="AT4g13270"
     CDS             18021..18599
                     /gene="AT4g13270"
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="CAB78369.1"
                     /db_xref="GI:7268030"
                     /translation="MASSKHEDYGIPYTPLPSSQPSQSVILLTPYRRHRPSLLRNLR
                     CSLLFTAVILLSAAVILLYPSDPDITVSRINLNHISVVDSHKIALDLSFSLITIKVRN
                     RDFPSLDYDSLVVSIGYRGRELGLVKSKGGHLKARDSSYIDATLELDGLEVHDVIYL
                     IGDLAKGVIPEDTIAQVQGDLGVLLFNIPIQV"
     exon            18021..18599
                     /gene="AT4g13270"
                     /number=1
     gene            19902..23486
                     /gene="AT4g13280"
     gene            complement(join(19902..20216,20354..20527,20683..20821,
                     21008..21229,21336..21714,22933..23209,23385..23486))
                     /gene="AT4g13280"
     CDS             complement(join(19902..20216,20354..20527,20683..20821,
                     21008..21229,21336..21714,22933..23209,23385..23486))
                     /gene="AT4g13280"
                     /note="similarity to predicted protein, Arabidopsis
                     thaliana"
                     /codon_start=1
                     /product="putative protein"
                     /protein_id="CAB78370.1"
                     /db_xref="GI:7268031"
                     /db_xref="GOA:Q9T0J9"
                     /db_xref="SPTREMBL:Q9T0J9"
                     /translation="MESQTTFKYESLAFTKLSHCQWTDYFLSVPIDESELDVITREID
                     ILKPEVMELLSSQGDDETSKRKVLLIQLLLSLGLAFHFENEIKNILEHAFRKIDDITG
                     DEKDLSTISIMPRVFRTYGHNLPSSVFKRFTGDDGKFQQSLTEDAKGILSLYEAAHLG
                     TTTDYILDEALKFTSSHLKSLLAGGTCRPHILRLIRNTLYLPQRWNMEAVIAREYISF
                     YEQEDHDKMLLRLAKLNFKLLQLHYIKELKSFIKWWMELGLTSKWPSQFRERIVEAW
                     LAGLMMYFEPQFSGGRVIAAKPNYLLTILDDACDHYFSIHELTRLVACVERWSPDGID
                     TLEDISRSVFKLMLDVPDDIGKGVRSEGSSYHLKEMLEELNTLVRANLDLVKWARGIQ
                     VPSFEBHVEVGGIALTSYATLMYSFVGMGETAGKEAYEWVFLYCQSDMSNGFAANAIN
                     YYMKQFVVTKEEAILECQRMIVDINKTINEBLLKTTSVPGRVLKQALNFGRLLELLYT
                     KSDDIYNCSEGLKEYIVTLLIDPIRL"
     exon            complement(19902..20216)
                     /gene="AT4g13280"
                     /number=1
     intron          complement(20217..20353)
                     /gene="AT4g13280"
                     /number=1
     exon            complement(20354..20527)
                     /gene="AT4g13280"
                     /number=2
     intron          complement(20528..20682)
                     /gene="AT4g13280"
                     /number=2
     exon            complement(20683..20821)
                     /gene="AT4g13280"
                     /number=3
     intron          complement(20822..21007)
                     /gene="AT4g13280"
                     /number=3
     exon            complement(21008..21229)
                     /gene="AT4g13280"
                     /number=4
     intron          complement(21230..21335)
                     /gene="AT4g13280"
                     /number=4
     exon            complement(21336..21714)
                     /gene="AT4g13280"
                     /number=5
     intron          complement(21715..22932)
                     /gene="AT4g13280"
                     /number=5
     exon            complement(22933..23209)
                     /gene="AT4g13280"
                     /number=6
     intron          complement(23210..23384)
                     /gene="AT4g13280"
                     /number=6
```

```
exon      complement(23385..23486)
          /gene="AT4g13280"
          /number=7

gene      27422..29411
          /gene="AT4g13290"

CDS       join(27422..27808,28045..28521,28605..28886,29085..29411)
          /gene="AT4g13290"
          /note="strong similarity to cytochrome P-450LXXIA1, Persea
          americana, M32885
          Contains Cytochrome P450 cysteine heme-iron ligand
          signature AA426-435
          contains EST gb:AI999004.1"
          /codon_start=1
          /product="cytochrome p450-like protein"
          /protein_id="CAB78371.1"
          /db_xref="GI:7268032"
          /db_xref="GOA:Q9T0K0"
          /db_xref="SWISS-PROT:Q9T0K0"
          /translation="MEIIVTLCLTTLLALLLLKSILKRTTNNLNLPPSPWRLPVIG
          NLHQLSLNTHRSLRSLRYGPLMLLHFGRTPVLIVSSADVAHDILKTYDVICANRPK
          TKVIDKILRGGRDVAFAPYGEYWKQMKSICIQNLLSNKMVRSYKKIREDEIKLMIEKV
          ENASSCSPPSPVNLSQLFMLTNDIICRAALGRKYSSKEDGIDVENIVRAFSALVGEF
          PIGEYIPSLSWIDKIRGQDHKMEEVDKRPDEFLERVVKEHEDANKDTRSDLVDTLLTI
          QSDKSALKLIIWDMFLAGTATSLSIEWAMTELMRNPKVMKKLQEIRSSSRQGLFVT
          EKEAEKMDYLQAVIKEALRLRPPAPLMVPRVFSEDVTLKGYNIPAGTQVIINAWAIQR
          DTTWGIDAEEFRPERHLDSILDFGQDFKFIPFGSGKRICPGIGFTSALIGVTLANI
          VKRFNWRMDVEPQRVQHDLTEATGLVVFRKFPLIAIPSSA"

exon      27422..27808
```

```
Query Match          44.3%;    Score 755.8;    DB 8;    Length 199634;
Best Local Similarity 71.7%;    Pred. No. 5.5e-124;
Matches 1036;    Conservative   0;    Mismatches 397;    Indels 12;    Gaps 3;

Qy    136 ACGCCTCGCCGCCGCCTCCCTTATCTGCGGCCATTCGCCGCGTGGACAGCAAGTTCG     195
Db 145005 ATGCCACGGCGCGGCCGCGGCCAACTACATCTGTGGCCAGTTAGGCGACGTCAACAACAAATTTA 145064

Qy    196 TCGACACGGCCGCTTCGCCGTCGACAAACACCTACCTCCTCTTTTCCGCCTACCTCGTTTTTT  255
Db 145065 TCGACACCGCTTTCGCTATAGACAACAACTTACCTTCTCTTTCTCCGCCCTACCTTGTCTCT 145124

Qy    256 CTATGCAGCTCGGCTTCGCCATGCTCGCGCGCCTCCGTCCGCGCCAAGAACACCATGA  315
Db 145125 CTATGCAGCTTGGCTTCGCTATGCTCGTGTGCCGGGTTCCGTGAGAGCCAAGAATACTATGA 145184

Qy    316 ACATCATGCTCACCAAGTCCTGGACGCGTCGCCGCGGCCGCCCTTCTCCACTACCTCTTCG  375
Db 145185 ACATCATGCTTACCAACGTCCTTGGACGCGTCGCTGCCGCCGGTGTCTCTTCATTTATCGTTTG 145244

Qy    376 GCTTCGCCTTCGCTTCGGCTCCCCCTCCAACGGCTTCATCGGTAAACATTTCTTCCGGCC  435
Db 145245 GCTACCGCCTTTGCCTTTGGATCTCCGTCCGTCCAATGGTTTCATCGGTAAACACTACTTTGGTC 145304

Qy    436 TCAAGACATCCCTTCATCCTCCTAGCACTACAGCTACTCCTCCTACCACAATGGGCCTTCG  495
Db 145305 TCAAAGACATCCCACGGCCTCTCGACTACTCGCTGACTACTCCAACTTTCTCTACCAAATGGGCCTTTG 145364

Qy    496 CCATCGCCGCGCCGCCGGCGATCACCAGGCGGAAGCATCGCCGAACCCACACAGTTCGTGGCCT  555
Db 145365 CAATCGCTGCGGCTGGAATCACCAAGTGGCTCGATGCTCGCGAACGGAACCACAGTTCGTGGCTT 145424

Qy    556 ATCTCATCTACTCCTCCTTCGTCACCGGCTTCGTCATCCCGGTGGTCTCCCACTGGTTCT  615
Db 145425 ACCTAATTCTATTCCTCTTTTGTTACCGGGTCGTCTCTCACTGGTTCT 145484

Qy    616 GGTCCCCAGACGCGTGGGCCTCGCCTT---TAAGATCACCGACCGGGTATTTCCACCG  672
Db 145485 GGTCAGTTGATGGATGGCGCCAGCCCGTTCCGTACCGATGGAGATTTGCTTTTCAGCACCG 145544

Qy    673 GCGTAATAGACTTCGCCGGTCCGGGTCCGGGCGTAGTCCACATGGTCGGCGGAATAGCCGGCCTAT  732
Db 145545 GAGCGATAGAGTTCCGTGGGTGTTGTTCATATGGTCGGAGGTTCGGAGGGTATCGCTGGACTCT 145604

Qy    733 GGGGAGCGGCTGATCGAAGGCCCAAGAATGGACGTTCGATCATCGAGGACGAGCTGTGG  792
Db 145605 GGGGTGCGCTCATCGAAGGCTCCACGACTTGCCGGTTCGATAAACGAGGGCCGTGCCATCG 145664

Qy    793 CCTTGCGAGGCCACAGCGCGTCCTTAGTAGTCCTGGGAACCTTCTTGCTTTGGTTCGGTT  852
Db 145665 CTCTTCGTCGTGGCCACTCGGCGTCACTTGTGTCCTTGGAACATTCCTCCTCTGGTTTGGAT 145724

Qy    853 GGTACGGATTTAACCCCGGTTCATTTAACAAAATCCTACTTACTTACGGTAACTCAGGAA  912
Db 145725 GGTACGGATTTAACCCCGGTTCCTTCAACAAGATCCTAGTCACGTACGAGA---CAGGCA 145781

Qy    913 ATTACTACGGTCAATGGAGCGGCGGTTGGCAGAACCGCGGTCACCACTACCCTAGCGGGGT  972
Db 145782 CATACAACGGCCAGTTGGAGCGGCGGTTCGGACCGGACACAGCTGTCACAACACGTTAGCTGCCT 145841

Qy    973 CAACAGCTGCCTTGACCACCAGCTATTCGGTAAACGGGTGAAAGCGTGAATCCGGTCACTGGAACGTGA 1032
Db 145842 GCACCGCGGGCGGCTGACAACCCTATTTGGGAAACGTCTACTCTCGGGACATTGGAACGTCA 145901

Qy   1033 CCGATGTCTGCAACGGCGTGTAGGCGGTTTCGCGCGCGGCGATAACAGCCGGTTGCTCCGTGG 1092
Db 145902 CTGATGTATGCAACGGCGCCTCCTCGGAGGGGTTTGCAGCCATAACTGGTGGCGTCGTCTGTCG 145961

Qy   1093 TTGAGCCATGGCAGCCATCGTATGGGGTTTGTTGTCTTCTATAGTATTAATAGCTTGCA 1152
Db 145962 TTGAGCCATGGCGTCGATCATCTGCGGCGGTTCGTGCGGGGGCCCTAGTCCTCCTCGGATGCA 146021

Qy   1153 ACAAATTAGCAGAGAAGGTTAAGTTCGACGATCCTCCTGGAGGCGGCGGCAGTTGCACGGTG 1212
Db 146022 ACAAGCTCGCTGAGGAGCTCAAATACGACGACCCCTTGAGGCAGCACAACTACACGGTG 146081

Qy   1213 GGTGTGGCACGTGGGGGGTGGATATTCACCGGCGTTGTTCGCAAAAAAGGAGTATGTGAAGG 1272
Db 146082 GTTCGGGTGTCGTGGGGACTAATATTCACGGGCTCTCTTCGCTCAAGAAAAGTACTTGAACC 146141

Qy   1273 AGGTTTACGGGTTG------GGGAGGGCGCACGGGTTGCTCATGGGGGGTGTGGGGAAGT 1326
Db 146142 AGATTTACGGCAACAAACCCGGAAGGCCACACGGTTTGTTATGGCGGTGGAGGAAAAC 146201

Qy   1327 TGCTTGGCGGCGCACGTGATTCAGATTCTGGTGATTGCTGGTGGTTAGTGCGACCATGG 1386
Db 146202 TACTTGGAGTCAGCTGATTCAGATCATTGTGATCACGGGTTGGGTAAGTGCGACCATGG 146261

Qy   1387 GACCCTTGTTTTGGGGGTTGAATAAACTGAAGCTGTTGAGGATTTCTTCAGAGGATGAGC 1446
Db 146262 GGACACTTTCTTCATCCTCAAGAAATGAAATTGTTGCGGATATCGTCCAGGATGAGA 146321

Qy   1447 TTGCGGGATGGACATGACTCGGCCATGGAGCTTTGCTTATGCTTATGAGGATGATGAGA 1506
Db 146322 TGGCCGGTATGGATATGACCAGGCACGGCGTGTTTGTTTGCTTATATGTACTTTGATGATGATG 146381

Qy   1507 CGCACAAGCATGGGATGGCAGTTGAGGAGGGTTGGGCCCAACGCGTCTTCCACACCCACCA 1566
Db 146382 AGTCTCACAAAGCCATTCAGCTTAGGAGAGTTGAGCCACGAGTTGAGCCACGATCTCCTTCTCCTTCTCGTG 146441

Qy   1567 CTGAT 1571
Db 146442 CTAAT 146446
```

```
REFERENCE
AUTHORS   1
TITLE     Sasaki,T., Matsumoto,T. and Yamamoto,K.
          Oryza sativa nipponbare(GA3) genomic DNA, chromosome 2, BAC
          clone:OJ1372_D06
JOURNAL   Published Only in Database (2001)
REFERENCE 2 (bases 1 to 69812)
AUTHORS   Sasaki,T., Matsumoto,T. and Yamamoto,K.
TITLE     Direct Submission
JOURNAL   Submitted (15-AUG-2001) Takuji Sasaki, National Institute of
          Agrobiological Sciences, Rice Genome Research Program; Kannondai
          2-1-2, Tsukuba, Ibaraki 305-8602, Japan
          (E-mail:tsasaki@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,
          Tel:81-298-38-7441, Fax:81-298-38-7468)
COMMENT   On Oct 30, 2003 this sequence version replaced gi:15208427.
          The orientation of the sequence is from -21M13 to M13rev of the BAC
          clone. The nucleotide sequence of this BAC clone was generated by
          combining Monsanto and RGP-Japan sequencing data.
FEATURES             Location/Qualifiers
     source          1..69812
                     /organism="Oryza sativa (japonica cultivar-group)"
                     /mol_type="genomic DNA"
                     /cultivar="Nipponbare"
                     /db_xref="taxon:39947"
                     /chromosome="2"
                     /clone="OJ1372_D06"
ORIGIN

Query Match            44.2%;  Score 753.8;  DB 8;  Length 69812;
Best Local Similarity  73.7%;  Pred. No. 1.2e-123;
Matches  988;  Conservative    0;  Mismatches  347;  Indels    6;  Gaps    2;

Qy     174 GCGCGCCTGGACAGCAAGTTCGTCGACACGGCCTTCGCCGTCGACAACACCTACCTCCTC    233
           ||||| ||||  || ||||||||||||| ||   ||||||| |  |  ||||| |||||||
Db   37437 GACTACCTGTGCAACCAGTTCGCGGACGACGTTCGGCGGCCGTGACTCGACGTACCTGCTC  37378

Qy     234 TTTTCCGCCTACCTCGTTTTTTCTATGCAGCTCGGCCTTCGCCATGCTCTCGCCGCCGTCC    293
           || ||| | |||||||| |  | |||||||||||||||||||| |||||| |||||||  ||
Db   37377 TTCTCGGCGTACCTCGTCGTTCGGCCATGCTCGGCCTTCGCCATGCTCGGCCGCCGGGTCC  37318

Qy     294 GTCCGCGCCAAGAACACCATGAACATCATGCTCACCAACGTCTGGACGTCTGGAGGCTCGGAGG    353
           |||||||||||||||||||||||||||||||||||||||| | |||||  |||||||||| |||
Db   37317 GTCCGCGCCAAGAACACCATGAACATCATGCTCACCAACGTGCTCGACGTGCTCGACGCCCGCCGCCGGC  37258

Qy     354 GGCCTCCTTCTACTACCTCTTCGGCCTTCGCCTTTCGCCTTTTCGGCTCCCCCCTCCAACGGCTTC    413
           ||||| |||||||||||| |||||| |||||||  |||||   |||   |||||||||||||||||
Db   37257 GCGCTCTTCTACTACCTCTTCGGCCTTCGCCTTCGGGGCGCGCCGGGGTCCAACGGCTAC  37198

Qy     414 ATCGGTAAACATTTCTTCGGGCCTCAAGGACATCCCTTCATCCTCCTACGACTACAGCTAC    473
           |||||| || |  | ||||| ||||||||| ||  | |||| | |||| |||| ||||||||
Db   37197 ATCGGGAAGCACTTCTTCGGGCCTCAAGGACAGTCCCACAGGTCGGCTTCGACTACAGCTTC  37138

Qy     474 TTCCTCTACCAATGGGCCTTCGCCATCGCCGCCGCCGCCGCGGAGAAGCATCGCC    533
           ||||| || || ||| |||||||||||||| |||| ||||||  ||||||||||||
Db   37137 TTCCTCTTCCAGTGGGCCTTCGCCTATCGCCGCCGGGGCTCCATCGCC  37078

Qy     534 GAAGGCACACAGTTCGTCGGCCTATCTCCTTCCTCCTCCTCGTCTATCGCTCGTCTAT    593
           ||||| ||||||||||||||| | |||||||||||||| ||| ||||||||||||||
Db   37077 GAGCGGCGACCCAGTTCGTGGCGTACCTCGTGGCGCCCTCCTCCACCGCGTTCGTGCTGTAC  37018

Qy     594 CCGGTGGTCTCCCACTGGTTCGTCCCCCAGACGGCTCTGGGCCCTCTGCCTTTAAGATCACC    653
           |||||||  |||||||||||||| | |  |||||||| || |||||||||| |  | |||||
Db   37017 CCGGTGGTGTCCCACTGGGATCGTGGTGCGGGCCCGACGGCGTTCCCCGGACGTCG  36958

Qy     654 G--ACCGGCTATTTTCCACGGCGTAATAGACTTCGCCGGTTCCGGCGTAGTCCACATG    710
           |   || |||  |||  ||||| |||||| |||||||| |||||||| |  | | ||||
Db   36957 GGGTCGTCGTCGCTTTGGGGTCCGGGGCTCATCGGCCTCGCGGGGGTCAGGGGTTGTCCACATG  36898

Qy     711 GTGCGGCGAATAGCCGGCCATGGGGAGCGGCGTGATCGAAGGCCCAAGAAATGGGACGTTTC    770
           ||| || | |||| ||| ||||||| ||||| |||| |||| || || |||| ||| ||| |
Db   36897 GTGGGCGCGGCGTCGGCCGGACCTCGTGGCGGCCCTCACCTCCGAGGGCCCCGCATTCGGCGGTTC  36838

Qy     771 GATCATGCAGGACGACGAGTCTGGCCTTGCGACGGCCACAGCCGCGGTGTCCTCCTAGTAGTCCTCGGGA    830
           |||| | || ||  |||||| |||||||||||| ||||||||||||||||||||||||||||||| ||||
Db   36837 GACCACGCAGGACGGCGAGTCTGGCCTTGCGACGGCCACAGCCGCGGTGTCCTCCTAGTAGTCCTCGGGC  36778

Qy     831 ACTTCTTGGTTCGGTTGGTACGGATTAACCCGGTTCATTTAACAAAATCCTA    890
           |||||||| ||| |||||| |||||| || || ||| |||||||||| ||||
Db   36777 AGCTTCCTTCGTGTGGTTCGGGTGGTACGGGTTAACCCGGTTCCTCACCATCTC  36718

Qy     891 CTTACTTACGGTAACTCAGGAAATTACTACCGTCAATGAGCAGGCGGGTTGGCAGAACCGCG    950
           || | || |||||||| ||||| ||| ||||||||| | || |||||| ||||||||||||
Db   36717 AAATCCTACGGCCCCGGGCTAGCATCCACGGCAGTGGTCGGCCAGTGTCGGGGACGCACCGCC  36658

Qy     951 GTCACCACTACCCTAGCGGGGTCAACAGCTGCCTTGACCACGCTATTCGGTAAACGGGTG    1010
           ||||||||||||| | |||| ||||||||| ||||  |||| |||||||| |||| |||
Db   36657 GTGACCACCACCCCTCGCCGGCAGCACGCGGCCGCTCCTCCTGCGGCCGTGTTCGCGGCG  36598

Qy    1011 ATATCCGGTCACTGGAACGTGACCGATGTCGCAACGGGCCGTGTTAGGCGGTTTCGCGGCG    1070
           | | ||||| ||||  ||||||||||||| || ||||| |||||| ||| ||||||| ||||
Db   36597 CAGACGGGCCACTGGAACGTGATGACGTGACGCTGCAACGGCCTCCTCGGCGGTTCGTCTCG  36538

Qy    1071 ATAACAGCCGGTTGCTCCGTTGGTTGAGCCATGGGCAGCCATCGTATGCGGTTTTGTTCGT    1130
           | |  |||||||||| | ||||| || ||||||||||||||||| | |||||||||| |||
Db   36537 ATCACCGCCGGTTGCTCCGTGCTGCCACCGTGGGCGGATCATCGCGGGTTCGGGTTCGTCTCG  36478

Qy    1131 TCTATAGTATTAATAGCTTGCAACAAATTAGCAGAGAAGGTTAAGTTCGACATCCTCTG    1190
           || ||||| || |||||||| ||||| | || |||||||| || ||||||||||| |||
Db   36477 GCGTGGGTGCTCATCGGCCTCAACGCGCTCGCGGCCTCAAGTTCGACGACACCCGTC  36418

Qy    1191 GAGGCGGCGCAGTTGCACGGTGGTGTGCACGTGCGGGGTGATATTCACCGGCGTTGTTC    1250
           |||||||| |||||||||||| ||||||||||| | ||||| ||| ||||||||| |||
Db   36417 GAGGCGGCGGCAGTTGCACGGCGGTGTGCACGCCGTGGGGGTGTCATCTTCACCGGCGCGTGTTC  36358

Qy    1251 GCAAAAAGGAGTATGTGAAGAGGTTACGG---GTTGGGGAGGGCGCACGGGTTGCTC    1307
           || || | |||  ||||| ||||||| ||   || || |||  || || |||| |||
Db   36357 GCGCGCAAGGAGTACGTGGACGATCTTCGGCGGCGGCGGGGCCCCGGCGCCGGTCGGGTC  36298

Qy    1308 ATGGGGGTGGTGTGGGAAGTTGCTGGCGGCCACGTGATTCAGATTCTGGTGATTGCTGGGG    1367
           ||||| ||||| ||||||||||||||| || |||| || |||||| ||| ||||| |||||
Db   36297 ATGGGGGCGGCGGCGGCCGGCGGCGGCCGCCACATAGTGGTCAAGTGGTCATCCTGGTCATCGGCGCG  36238

Qy    1368 TGGGTTAGTGGACCATGGACCCTTGTTTGGGGTTGAATAAAACTGAAGCTGTTGAGG    1427
           |||||| |||| |  ||  | |||| || ||| ||| || |||  ||| ||||| |||
Db   36237 TGGGTGAGCTTCACCATGGCGCCGCTGTTCCTGTGCTCAACAAGCTGGGCGTTGCGCGC  36178

Qy    1428 AATTTCTTCAGAGGATGAGCTTGCCGGGGATGGACAATGACTCGCCATGGAGGCTTTGCTTAT    1487
           |||||||| |||| ||||||| ||| || | || |||||||| |||||||||| |||| ||||
Db   36177 ATCTCGGCCGCGAGGACGAGATGGCCGCCATGGACGAGACCACCGCCACGGGGCCCACGGGTAC  36118

Qy    1488 GCTTATGAGGATGATGAGACG    1508
           || |||||||||||||| |||
Db   36117 GGGTACCACGACGACGACGCG  36097
```

Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y.,
Kurosaki,T., Kodama,T., Masuda,H., Kobayashi,M., Xie,Q., Lu,M.,
Narikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Niikura,J.,
Ikeda,J., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J.,
Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN:,
Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S.,
Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M.,
Kagawa,I., Kondo,S., Konno,H., Miyazaki,A., Osato,N., Ota,Y.,
Saito,R., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
Yoshino,M. and Hayashizaki,Y.

TITLE   Collection, mapping, and annotation of over 28,000 cDNA clones from
          japonica rice
JOURNAL   Science 301 (5631), 376-379 (2003)
MEDLINE   22752273
PUBMED   12869764
REFERENCE 2 (bases 1 to 1853)
AUTHORS   Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K.,
          Fujimura,T., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W.,
          Hayashida,K., Hayashizaki,Y., Hayatsu,N., Hiramoto,K., Hiraoka,T.,
          Hori,F., Hotta,I., Iida,J., Iida,Y., Ikeda,R., Imamura,K.,
          Imotani,K., Ishibiki,J., Ishii,Y., Ishikawa,M., Itoh,M., Kagawa,I.,
          Kanagawa,S., Katoh,H., Kawagashira,N., Kawai,J., Kawamata,M.,
          Kikuchi,S., Kishikawa-Hirozane,T., Kishimoto,N., Kobayashi,M.,
          Kodama,T., Kojima,K., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
          Koya,S., Kurihara,C., Kurosaki,T., Kusumegi,T., Li,C., Lu,M.,
          Masuda,H., Matsubara,K., Matsuyama,T., Miura,J., Miyazaki,A.,
          Mizuno,K., Murakami,K., Murata,M., Nagata,T., Nakamura,M.,
          Namiki,T., Narikawa,R., Niikura,J., Nishi,K., Nomura,K.,
          Numasaki,R., Ohneda,E., Ohno,M., Ohtsuki,S., Oka,M., Ooka,H.,
          Osato,N., Ota,Y., Otomo,Y., Ryu,R., Saitoh,H., Sakai,C., Sakai,K.,
          Sakazume,N., Sano,H., Sasaki,D., Sato,K., Sato,K., Shibata,K.,
          Shinagawa,A., Shiraki,T., Shishiki,T., Sogabe,Y., Sugano,S.,
          Sugiyama,A., Suzuki,K., Suzuki,Y., Tagami,M., Tagami-Takeda,Y.,
          Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A.,
          Toya,T., Tsunoda,Y., Ueda,M., Waki,K., Xie,Q., Yahagi,W.,
          Yamada,H., Yamamoto,M., Yasunishi,A., Yazaki,J., Yokomizo,S. and
          Yoshimura,A.
TITLE   Direct Submission
JOURNAL   Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of
          Agrobiological Sciences, Department of Molecular Genetics, Head of
          Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki
          305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp,
          Tel:81-29-838-7007, Fax:81-29-838-7007)
COMMENT   This clone is one of the 28K full-length cDNA clones from japonica
          rice.
          URL : http://cdna01.dna.affrc.go.jp/cDNA/
          NIAS Rice Full-Length cDNA Project Team: Kikuchi,S., Satoh,K.,
          Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J.,
          Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T.,
          Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T. and
          Yamamoto,M.
          FAIS Genome Sequencing & Analysis Group: Otomo,Y., Iida,Y.,
          Fujimura,T., Ikeda,R., Ishibiki,J., Kawamata,M., Kobayashi,M.,
          Kodama,T., Kurosaki,T., Kusumegi,T., Lu,M., Masuda,H., Miura,J.,
          Mizuno,K., Narikawa,R., Niikura,J., Oka,M., Ryu,R., Sugano,S.,
          Yoshimura,A., Matsubara,K. and Murakami,K.
          Genome Exploration Research Group in Riken Genomic Sciences Center
          and Genome Science Laboratory in Riken: Adachi,J., Aizawa,K.,
          Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hanagaki,T.,
          Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K.,
          Hiraoka,T., Hori,F., Iida,J., Imamura,K., Imotani,K., Ishii,Y.,
          Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawai,J.,
          Kishikawa-Hirozane,T., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
          Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
          Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Ohno,M., Osato,N.,
          Ota,Y., Saitoh,H., Sakai,C., Sakai,K., Sakazune,N., Sano,H.,
          Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
          Sogabe,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F.,
          Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Waki,K.,
          Yasunishi,A. and Hayashizaki,Y.
FEATURES   Location/Qualifiers
     source   1..1853

/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="002-125-B07"

ORIGIN

Query Match   44.1%;   Score 752.2;   DB 8;   Length 1853;
Best Local Similarity 73.6%; Pred. No. 2.4e-123;
Matches 987; Conservative 0; Mismatches 348; Indels 6; Gaps 2;

```
Qy    174 GCCGCCGTGGACGACGAAGTTCGTCGACAACGGCCTTCGCCGTCGACAACACCTACCTCCTC 233
              |||| ||||||  || ||||| ||||| || ||||| ||||||  ||||| |||  |||
Db    132 GACTACCTGTGCAACCAGTTCGGCGACGACGAGTCGGCGTCGGACTCGACATACCTGCTC 191

Qy    234 TTTTCCGCCTACCTCGTTTTTTCTGGCCGTCGCCATGCTCGTGCCGCCGGCTCC 293
              ||| | |||||||||| || | ||||||||||||||||||||||||||||| |
Db    192 TTCTCGGCGTACCTCGTCGTCGTTGGCCGTCGCCATGCTCGTGCCGCCGGTCC 251

Qy    294 GTCCGCGCCAAGAACACCATGAACATCATGCTCACCAACGTCCTGGACGCTGCCGCCGCCGGC 353
              |||||  |||||||||||||||||| | ||||| |||||||| || |||||| ||||||||||
Db    252 GTCCGGCGTCAAGAACACCATGAACACCATGCTTACCAACGTGCTCGACGCGCCGCCGCCGGC 311

Qy    354 GGCCTCCTCTACTACCTCTTCGGCTTCGCCTTCGGCTTCGGCTCCCCCCCCTCCAACGGCTTC 413
              | || ||||||||||| ||||||||||||||||  ||| | |  ||   ||||||||| |||
Db    312 GCGCTCTTCTACTACCCTCTTCGGCTTCGCCTTCGGGCTTCTCGGGCGCCGTCCAACGGCTTC 371

Qy    414 ATCGGTAAACATTTCTTCGGCCTCAAGGACATCCTTCAGCATCCTCTACGACTACAGCTAC 473
              |||||| | || |  |||| | ||||||||  |  ||| ||| || ||||| |||||||||
Db    372 ATCGGGAACCACTTCTTCGGCCTCAAGCAGTCCCACAGGTCGGCTTCGACTACAGCTTC 431

Qy    474 TTCCTCTACCAATGGCCTTCGCCGCCTCTGGCCTCTCATCTCCTCCTTCTCATCGGCGCTCTAT 533
              |||| |  |||||| ||||||||||||| | ||   |  |||| || | ||| ||| |||
Db    432 TTCCTTCCCAGTGGGCCTTCGCCGCCTTGCCCCGCGGGCATCACGCGCCGGCTCCATGCGCC 491

Qy    534 GAAGGCACACAGTTCGTGGCGTTCGTCGTGGGCTACCTCATCTCCTCCTTCTCATCTCGTCTAT 593
              ||| |||| |||||||| |||  |||| |||||| ||||| || ||||| |||||| | |||
Db    492 GAGCGGACCCCAGTTCGTGTGGGTACCTCATCTACTCCGCCTTCCTCACCGGCGTTCGTCTAC 551

Qy    594 CCGGTGGTCTCCCACTGGTTCTGGTCCCAGACGGCGTGGGCCTGGCCCTTTAAGATCACC 653
              ||||||||| || || || ||  ||||||||||||||||| || |  || ||| ||||||
Db    552 CCGGTGGTCGTCCCACTGGATCTGGTCCCGGCGGCGTGGGCCTTCGGCTTCCCCGGACGTCG 611

Qy    654 G---ACCGGCTATTTCCACCGGCGTAATAGACTTCGCCGGTTCCGGCGTAGTCCACATG 710
                  ||| |||||||||  | |||| | || ||||||| ||| || ||||  | ||||||
Db    612 GGGTCGCTGTCTCTTCGGGTCCGGCGCATCGACTTCGCCGGTCAGGGGTTGTCCACATG 671

Qy    711 GTCGGCGGAATAGCCGGCTCATGGGGAGCCGCTGATCGAAGGCCCAAGAATGGACGTTTC 770
              |||||||||||||| | || ||||| | |||| |  ||  ||||||||||||| |||||
Db    672 GTGGGCGGCGTGGCCGGCGACTCTGGGGGCGCCCCATCGAGGGCGCCCATTGGGCGGGTC 731

Qy    771 GATCATGCAGGACGACGCTGTGGCCTTGGAGGGCCACAGCGGCGTCCTTAGTAGTCCTGGGA 830
              |||| |||||||||||| | | | ||||| | ||||||| |||||| | ||||  ||| ||
Db    732 GACCACGCCGGCGGCGCTGCGGGGCTTGGGGCGCCACAGCGGCGTCGTCGTCGTCGTCGGC 791

Qy    831 ACCTTCTTGCTTCGTGTTCGGTTCGGTTGTACGGATTAACCCGGTTCATTTAACAAAATCCTA 890
              |||||||||| || |  ||||| ||||| |  | |||| |||||| ||  | |||| | |||
Db    792 AGCTTCCTTCTGTTCGGTTCGGTTACGGTGTACGGGTTTAACCCCGGCTCGTCGTTCCTCACCATCCTC 851

Qy    891 CTTACTTACGGTAACTCAGGAAATTACTACGGTCAATGAGCGCGGTTGGCAGAACCGCG 950
              ||||| | |||||| | |||||||  ||||  || ||| | ||||  || | ||||  ||
Db    852 AAATCCTACGGCCCCGCCGGTAGCATCCACGGGACAGTGTGGGCGGCACCGCC 911

Qy    951 GTCACCACTACCCTAGCGGGTCAACAGCTGCCTTGACCACGCTATTCGGTAAACGGGTG 1010
              ||||| | | |||    | ||  ||||||  ||||| |||| |||| |||| ||| |||
Db    912 GTGACCACCACCCTTGCCGGCAGCAGCGCTCACGACGCGCTTCGGGAAGAGGCTC 971

Qy   1011 ATATCCGGTCACTGGAACGTGACCGATGTCTGCAACGGGTCGTTGCGCGGCG 1070
              | |||||| ||||| ||| | || || | |  ||||||| | ||  | |||
Db    972 CAGAGGGGGCACTGGAACGTGAGCGTGACGTGCAAGCGCTCCTCGCGGCG 1031

Qy   1071 ATAACAGCCGGTTGCTCCGTGGTTGAGCCATGGCAGCCATCGTATGGCGGTTTTGTTGCT 1130
              | ||||| |||| || |||  |  || |||||||  | || | ||| ||||| ||||||
```

```
Db    1032 ATCACGCGCCGGTTGCTCCGTCGTCGACCCGTGGGCCGCGATCATCTGCGGGTTCGTCTCG 1091

Qy    1131 TCTATAGTATTAATAGCTTGCAACAAATTAGAGAGAGAGGTTAAGTTCGACGATCCTCTG 1190
                                     |||||||||||||||
Db    1092 GCGTGGGTGCTCATCGGCCTCAACGCGCTGGCGGAGGCTCAAGTTCGACGACCCGCTC 1151

Qy    1191 GAGGCGGCGCAGTTGCACGGTGGGTGTGGCACGTGGGGGGTGATAATTCACGGCGTTGTTC 1250
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1152 GAGGCGGCGCAGCTGCACGGCGGGGGTGCGGCGTGGGGGGTCATCTTCACGGCGCTGTC 1211

Qy    1251 GCAAAAAGAGTATGTGAAGGAGGTTTACCG---GTTGGGGAGGGCGCACGGGTTGCTC 1307
                   |||||||||||||||||||||||||    ||||||||||||||||
Db    1212 GCGCGCAAGGAGTACGTGGACCAGATCTTCGGCCAGCCGGCGCCCGTACGGGCTGTTC 1271

Qy    1308 ATGGGGGGTGTGGGAAGTTGCTGGCGGGCGCACGTGATTCAGATTCGGTGATTGCTGGG 1367
           |||||||| ||||||||||||||||||||||||||||||||
Db    1272 ATGGGGCGGCGGCCGGCCGGCGCCACACATAGTGGTCATCGGGTCATCGCGGCG 1331

Qy    1368 TGGGTTAGTGCGACCATGGGACCCTTGTTTTGGGGGGTTGAATAAACTGAAGCTGTTGAGG 1427
           ||||||| |||||||||||||||||||||||||||||||| |||||
Db    1332 TGGGTGAGCTTCACCATGGCGCGCGCTGTTCCGTGCTCAACAAGCTGGGCTTGCTGCGC 1391

Qy    1428 ATTTCTTCAGAGGATGAGCTTGCGGGGGATGGACAATGACTCGGCCATGGGAGGCTTTGCTTAT 1487
                |||||||||||||| |||||||||||||||||||| ||||||
Db    1392 ATCTCGGCCGAGGACGAGATGGCCGGCATGGACCAGACGACGCGCCACGCGGGGTTCGCGTAC 1451

Qy    1488 GCTTATGGAGGATGATGAGACG 1508
           || || |||| ||||||
Db    1452 GGGTACCACGACGACGACGCGCG 1472


RESULT 13
AP004053/c
LOCUS       AP004053    128913 bp    DNA    linear    HTG 21-MAR-2002
DEFINITION  Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 2 clone
            OJ1234_B11, *** SEQUENCING IN PROGRESS ***.
ACCESSION   AP004053
VERSION     AP004053.1  GI:15208421
KEYWORDS    HTG; HTGS_PHASE2.
SOURCE      Oryza sativa (japonica cultivar-group)
ORGANISM    Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1
  AUTHORS   Sasaki,T., Matsumoto,T. and Yamamoto,K.
  TITLE     Oryza sativa nipponbare(GA3) genomic DNA, chromosome 2, BAC
            clone:OJ1234_B11
  JOURNAL   Published Only in Database (2001)
REFERENCE   2  (bases 1 to 128913)
  AUTHORS   Sasaki,T., Matsumoto,T. and Yamamoto,K.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-AUG-2001) Takuji Sasaki, National Institute of
            Agrobiological Resources, Rice Genome Research Program; Kannondai
            2-1-2, Tsukuba, Ibaraki 305-8602, Japan
            (E-mail:tsasaki@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,
            Tel:81-298-38-7441, Fax:81-298-38-7468)
COMMENT     The nucleotide sequence of this BAC clone was generated by
            combining Monsanto and RGP-Japan sequencing data.
            NOTE: It currently consists of 1 contigs. Gaps between the contigs
            are represented as runs of N. The order of the pieces is believed
            to be correct as given, however the sizes of the gaps between them
            are based on estimates that have provided by the submitter. This
            sequence will be replaced by the finished sequence as soon as it is
            available and the accession number will be preserved.
            * NOTE: This is a 'working draft' sequence.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
FEATURES    Location/Qualifiers
     source 1..128913.
            /organism="Oryza sativa (japonica cultivar-group)"
            /mol_type="genomic DNA"
            /cultivar="Nipponbare"
            /db_xref="taxon:39947"
            /chromosome="2"
            /clone="OJ1234_B11"


ORIGIN
            Query Match    43.4%;    Score 739.6;   DB 2;    Length 128913;
            Best Local Similarity  73.5%;   Pred. No. 4.1e-121;
            Matches 986; Conservative    0; Mismatches 349;   Indels   7; Gaps   3;

Qy    174 GCCGGCCTGGACAGCAAGTTGTGACACGGCCTTCGCGTGCTGACAACACTACCTCCTC 233
          ||||||||||||||||| || |||||||||||||| |||||||||||||||||||
Db 118858 GACTACCTGTGCAACCAGTTCGCGGACACGACTCGGCGCTGTGCGACGTACCTGCTC 118799

Qy    234 TTTTCGCCTCACCTCGTTGTTTTTCTATGCAGCTGCGGCTTCGCTTCGCGGCCGGCTCC 293
          |||| ||||||| ||||||| |||||||||||||||||||||||||||||
Db 118798 TTCTGGCGTACCTCGTTCGCCATGCCATGCCTTCGCGGCTTCGCGGCCGGGTCC 118739

Qy    294 GTCCGCGCCAAGAACACCATGAACATCATGGTCACCAAGTCTCTGGACGCTGCCGCCGGC 353
          |||||||||||||| || ||||||||||||||||||||||||| |||||||||||||
Db 118738 GTCCGCGCCAAGAACAACCATGAACATCATGGTCACCAACGTGCTGCGACCGCCGGCCGGC 118679

Qy    354 GGCCTCTTCTACTACCTCTTCGGCTTCGGCTTCCTCAAGGACATCCCTTCATCTCCTACAGACTACAGCTAC 413
          |||||||||| ||||| ||||||||||| ||||| |||||||||||||||||||||
Db 118678 GCGCCTTCTACTACCTCTTCGGCTTCGGCCTTCGGCCCGTCCAAGCGGTTCAACGGCTTC 118619

Qy    414 ATCGGTAAACATTTCTTCGGCCTCAAGGACATCCCTTCATCTCCTACAGACTACAGCTAC 473
          |||||||| |||||||||||||||||||||||||||||||||||||||||
Db 118618 ATCGGGAAGCACTTTTTCGGCCTCAAGGACATCCGCCCACACGTCGGCTTCGACTACAGCTTC 118559

Qy    474 TTCCTCACCAATGG---CCTTCGCCATCGCCCGCGCCGGCATCACCAGCGGAAGCATCGC 532
          |||||||||||    |||||||||||||||||||||||||||||| |||||||||||||
Db 118558 TTCCTTCCAGTGGGCCCCTTCGCCATCGCCCGGGCCATCACGTCCGGCTTCGGCTC CCATCGC 118499

Qy    533 CGAACGCACACAGTTCGTGGCCTATCTCATCTCATTCACTCCTCCTCACCGGCTTCGTCTA 592
          ||||||||||||||||||||||||||||||| || |||||||| |||||||||||||||
Db 118498 CGAGCGGACCCAGTTCGTGGCGTACCTCATTTACTCGGCCCTTCTCACCGGCCTTCGTCTA 118439

Qy    593 TCCGGTGGTCTCCACTGGTTCTCGGTCCCCAGACGGCTGGGGTCCTGGGCCTTTAAGATCAC 652
          ||||||||||||||||||| ||| ||||||||||| |||||||||||||||||||||
Db 118438 CCCGGTGGTGTCCACTGGATCTCGGTCGCCGACGGGTGGGCCTTCGGGCCTTCCCGGACGTC 118379

Qy    653 CG---ACCGGCTATTTCCACGCGGCTAATAGACTTCGCCGGTTCCGGGTAGTCCACAT 709
               ||||||||||||||||||||| ||| ||||||||||||||||
Db 118378 GGGGGTCGCTGCTCTTCGGGTCCGGGCTCATCGACTTCGCCGGTGCTCGGGGGTGTCCACAT 118319

Qy    710 GGTCGGCGGAATAGCCGGCCGGCCATGGGAGGCGAGGGCTGATCGGAAGGCCCAAGAATGGGACGTTT 769
          |||||||||||| ||| |||||||||| |||||||||||||||||||||||||||
Db 118318 GGTGGGCGGCGGTGGGCGGAGCTGTGGGGGCGCCCTCATCGAGGGGGCCCCCATTGGGCGGTT 118259

Qy    770 CGATCATGCGAGGAGACGAGCTGTGGCCTTGCGAGGCCACAGCGGCGTCCAGCGGCGTCGTGCTCCTGGG 829
          ||||||| |||||||||||||| |||| ||||||||||||||||||||||||||
Db 118258 CGACCACGCCGGCCGGCTGGATCTGGTGCTGGCGGCCACACAGCGCCGTCGGCGTCGTGCTCGG 118199

Qy    830 AACCTTCTTGCTTTGGTTCGTTGGTTACGGATTTAACCCGTTCATTAACAAAATCCT 889
          |||||||| ||| |  |||||  ||| |||||||| |||||||||||
Db 118198 CAGCTTCCTTCGTTTCGTGGTTCGTGGTGGTGGTTAACCCCGGCTGGTTCCTCACCATCCT 118139

Qy    890 ACTTACTTACGGTAACTCAGGAATTACTACGGTCAATGGACGGCCGGTTGGCCAGAACCGC 949
          |||||| ||| |||||||||||||||||||||||||||||||||||||||
Db 118138 CAAATCCTACGGCCCCGCCGTAGCATCCACGGGCAGTGGTGTCGGCGGGGACGACCGC 118079

Qy    950 GGTCACCACTACCCTAGCGGGGTCAACAGTGCCTTGACCAGTCGGTAAACGGGT 1009
          |||||||| ||||||||||||||||| |||||||||||||||||||||||
Db 118078 CGTGACCACCACCTCGCCGCGCAGCACCAGCGGCGCCTCAGACGCCTTCGGGAAGAGGCT 118019

Qy   1010 GATATCCGGTCACTGAACGTGACCGATGTCGCAACGGGCTGTGTTAGGCGGGTTTCGCGGGC 1069
          |||| ||||| ||| |||||||| |||||||||||||||||||| |||||||||||
Db 118018 CCAGACGGGGCACTGGAACGTGAGATCGGACGGCTGCAACGTTCCTCGCGGCTTCGCGGGC 117959

Qy   1070 GATAACAGCCGGTTGCTCCGTGGTTGAGCCATGGCAGCCATCGTATGCGTTTGTTGC 1129
          |||| |||||||||||||||||||||||||||||||||||||||||
Db 117958 GATCACCCGCGGTTGCTCCGTCGACCCGTGGGGCCGGATCATCTGCGGGTTCGTCTC 117899
```

```
Qy   1130 TTCTATAGTATTAATAGCTTGCAACAAATTAGCAGAGAAGGTTAAGTTCGACGATCCTCT 1189
               ||||   |||||||    |||||| |  |||||
Db 117898 GGCGTGGGTGCTCATCGGCCTCAACGCGCTGGCGGCCGAGGCTCAAGTTCGACGACCCGCT 117839

Qy   1190 GGAGGCGGCGCCAGTTGCACGGTGGGTGGGCACGTGGGGGGTGATATTCACGGCGTTGTT 1249
               |||||||| ||||| |||| ||||| |||||||||||| |||||| ||| |||| ||||
Db 117838 CGAGGCGGCGCCAGCTGCACGGCGCGTGCGGCCGTGGGGGGGTCATCTTCACGGCGCTGTT 117779

Qy   1250 CGCAAAAAGGAGTATGTGAAGGAGGTTACGG--GTTGGGGAGGGCGCACGGGTTGCT 1306
               |||    ||    | ||| |||||  |  |||||||  ||| | ||||||||| ||
Db 117778 CGCGCCGCAAGGAGTACGTGGGACCAGATCTTCGGCCAGCCAGGCCGCCCGTACGGGCGTT 117719

Qy   1307 CATGGGGGGTGGGTGGGAAGTTGCTGGCGGCGCCGCACGTGATTCGTGTGATTGCTGG 1366
               ||||||| | ||||  |||||||| |||||||||||||| | || |  |||||||||||
Db 117718 CATGGGCGCGCGGCGCCCGGCCGGCTGCGGGGGGCCACATAGTGGTCATCCTGGTCATCGGCGGC 117659

Qy   1367 GTGGGTTAGTGCGACCATGGGGACCCCTGTTTTGGGGGTTGAATAAACTGAAGCTGTTGAG 1426
               ||||   || | |||||||||| ||   ||  ||||||||||| ||  |||||  ||||||
Db 117658 GTGGGTGAGCTTCACCATGGCGCCGCCGTTCCTGGTGCTCACAAGCTGGCGTTGCTGCG 117599

Qy   1427 GATTTCTTCAGAGATGAGCTTGCGGGATGGACATGACCTCGCCATGGAGGCTTTGCTTA 1486
               |||||||| |||| | ||||||||| ||| ||||| || |||||| ||||||| |||| ||
Db 117598 CATCTCGGCCGAGGACGAGATGGCCGCCATGGACCAGAGCGCGCCACGGCGGGGTTCGGTA 117539

Qy   1487 TGCTTATGAGGATGATGAGACG 1508
               |||| |  |||| ||||| |||
Db 117538 CGCGTACCACGACGACGACGCG 117517


RESULT 14
AK073718
LOCUS       AK073718                2044 bp    mRNA     linear   PLN 24-JUL-2003
DEFINITION  Oryza sativa (japonica cultivar-group) cDNA clone:J033060L01, full
            insert sequence.
ACCESSION   AK073718
VERSION     AK073718.1  GI:32983741
KEYWORDS    FLI_CDNA; CAP trapper.
SOURCE      Oryza sativa (japonica cultivar-group)
  ORGANISM  Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1
  AUTHORS   The Rice Full-Length cDNA Consortium, National Institute of
            Agrobiological Sciences Rice Full-Length cDNA Project Team:,
            Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K.,
            Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I.,
            Kojima,K., Namiki,T., Yazaki,J., Ishikawa,M., Suzuki,Y., Li,C.,
            Ohtsuki,K., Shishiki,T., Foundation of Advancement of International
            Science Genome Sequencing & Analysis Group:, Otomo,Y., Murakami,K.,
            Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y.,
            Narikawa,R., Sugiyama,A., Mizuno,K., Kobayashi,M., Xie,Q., Lu,M.,
            Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Shiraki,J.,
            Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN:,
            Hara,A., Hashidume,W., Imotani,K., Imotani,K., Ishii,Y., Itoh,M.,
            Kagawa,I., Kondo,S., Konno,H., Miyazaki,A., Osato,N., Ota,Y.,
            Saito,R., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
            Yoshino,M. and Hayashizaki,Y.
  TITLE     Collection, mapping, and annotation of over 28,000 cDNA clones from
            japonica rice
  JOURNAL   Science 301 (5631), 376-379 (2003)
   MEDLINE   22752273
   PUBMED    12869764
REFERENCE   2  (bases 1 to 2044)
  AUTHORS   Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K.,
            Fujimura,T., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W.,
            Hayashida,K., Hayashizaki,Y., Hayatsu,N., Hiramoto,T., Hiraoka,T.,
            Hori,F., Hotta,I., Iida,J., Iida,Y., Ikeda,R., Imamura,K.,
            Imotani,K., Ishibiki,J., Ishii,Y., Ishikawa,M., Itoh,M., Kagawa,I.,
            Kanagawa,S., Katoh,H., Kawagashira,N., Kawai,J., Kawamata,M.,
```

Kikuchi,S., Kishikawa-Hirozane,T., Kishimoto,N., Kobayashi,M.,
Kodama,T., Kojima,K., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Kurosaki,T., Kusumegi,T., Li,C., Lu,M.,
Masuda,H., Matsubara,K., Matsuyama,T., Miura,J., Miyazaki,A.,
Mizuno,K., Murakami,K., Murata,M., Nagata,T., Nakamura,M.,
Namiki,T., Narikawa,R., Niikura,J., Nishi,K., Nomura,K.,
Numasaki,R., Ohneda,B., Ohno,M., Ohtsuki,K., Oka,M., Ooka,H.,
Osato,N., Ota,Y., Otomo,Y., Ryu,R., Saitoh,H., Sakai,C., Sakai,K.,
Sakazume,N., Sano,H., Sasaki,D., Sato,K., Satoh,H., Shibata,K.,
Shinagawa,A., Shiraki,T., Shishiki,T., Sogabe,Y., Sugano,S.,
Sugiyama,A., Suzuki,K., Suzuki,Y., Tagami,M., Tagami-Takeda,Y.,
Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A.,
Toya,T., Tsunoda,Y., Ueda,M., Waki,K., Xie,Q., Yahagi,W.,
Yamada,H., Yamamoto,M., Yasunishi,A., Yazaki,J., Yokomizo,S. and
Yoshimura,A.
TITLE   Direct Submission
JOURNAL   Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of
          Agrobiological Sciences, Department of Molecular Genetics, Head of
          Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki
          305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp,
          Tel:81-29-838-7007,Fax:81-29-838-7007)
COMMENT   This clone is one of the 28K full-length cDNA clones from japonica
          rice.
          URL : http://cdna01.dna.affrc.go.jp/cDNA/
          NIAS Rice Full-Length cDNA Project Team: Kikuchi,S., Satoh,K.,
          Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J.,
          Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T.,
          Ohneda,B., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T., and
          Yamamoto,M.
          FAIS Genome Sequencing & Analysis Group: Otomo,Y., Iida,Y.,
          Fujimura,T., Ikeda,R., Ishibiki,J., Kawamata,M., Kobayashi,M.,
          Kodama,T., Kurosaki,T., Kusumegi,T., Lu,M., Masuda,H., Miura,J.,
          Mizuno,K., Narikawa,R., Niikura,J., Oka,M., Ryu,R., Sugano,S.,
          Sugiyama,A., Suzuki,Y., Tsunoda,Y., Ueda,M., Xie,Q., Yokomizo,S.,
          Yoshimura,A., Matsubara,K. and Murakami,K.
          Genome Exploration Research Group in Riken Genomic Sciences Center
          and Genome Science Laboratory in Riken: Adachi,J., Aizawa,K.,
          Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hanagaki,T.,
          Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K.,
          Hiraoka,T., Hori,F., Iida,J., Imamura,K., Imotani,K., Ishii,Y.,
          Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawai,J.,
          Kishikawa-Hirozane,T., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
          Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
          Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Ohno,M., Osato,N.,
          Ota,Y., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H.,
          Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
          Sogabe,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F.,
          Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Waki,K.,
          Yasunishi,A. and Hayashizaki,Y.
FEATURES   Location/Qualifiers
     source   1..2044
                     /organism="Oryza sativa (japonica cultivar-group)"
                     /mol_type="mRNA"
                     /cultivar="Nipponbare"
                     /db_xref="taxon:39947"
                     /clone="J033060L01"
ORIGIN
Query Match   43.3%; Score 738.2; DB 8; Length 2044;
Best Local Similarity   72.4%; Pred. No. 7.2e-121;
Matches 972; Conservative 0; Mismatches 368; Indels 3; Gaps 1;

```
Qy    174 GCCGCCGGTGGACAGCAGCAAGTTCGTCGACACGCCGCCTTCGCCGTCGACAACACCTACCTCCTC 233
              ||  ||||||| ||||||| |||| |||| ||| ||||||| ||||||| || ||||| ||||
Db    145 GACTACCTGTGCAACCGGTTCGCCGACACGACGTCGGCGGTGGACGCGACGTACCTGCTC 204

Qy    234 TTTTCCGCCTACCTCGTTTTTCTCTATGCAGCTCAGCAGCGCCATGCTCCTCGCCGGCTCC 293
              |||||||| |  |||| | |  |  || || |||| || |||||||||| || |||  ||||
Db    205 TTCTCGGCGTACCTCGTGTTCGCCCATGCAGCCATCATGCTCCTCGGATGCTCCGCGATGGTCTCGCCGGCTCC 264

Qy    294 GTCCGGCGCAAGAACACCATGAACATCATGCTCACCAAGCTCCTGACGCGTGCGCCGGC 353
              ||||||||||||||||||||||||||||||||| ||||||||||| ||||| ||||| |
Db    265 GTGCGGGCCAAGAACACCATGAACATCATGCTCACCAAGCTGCTCACCACGCTGCGCCGGGGG 324
```

```
Qy  1431 TCTTCAGAGGATGAGCTTGCGGGATGGACATGACATGACTCGCCATGGAGCTTTGCTTATGCT 1490
         |  |||||| ||||  ||   |||  ||| |||| ||  | |||| ||||||| ||   ||
Db  1405 TCCGCCGAGGACGAGACGTCCGGCCATGGACCATGGACCTGACACGGCACGGCGGGTTCGCGTACGTC 1464

Qy  1491 TATGAGGATGATGAGACGAGACGACAA 1513
         || | || || ||| ||||| ||| |
Db  1465 TACCACGACGAGGACGACGACGACGA 1487
```

RESULT 15
AF289477

```
LOCUS       AF289477               4123 bp    DNA     linear   PLN 20-SEP-2001
DEFINITION  Oryza sativa ammonium transporter 1-1 (Osamt1-1) gene, complete
            cds.
ACCESSION   AF289477
VERSION     AF289477.1  GI:15705367
KEYWORDS    .
SOURCE      Oryza sativa
  ORGANISM  Oryza sativa
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 4123)
  AUTHORS   Hoque,M.S., Masle,J., Udvardi,M.K. and Upadhyaya,N.M.
  TITLE     Cloning and characterization of three Ammonium transporter genes
            from rice
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 4123)
  AUTHORS   Hoque,M.S., Masle,J., Udvardi,M.K. and Upadhyaya,N.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-JUL-2000) Gene Expression and Plant Improvement,
            CSIRO Plant Industry, Cnr. of Barry Drive and Clunies Ross Street,
            Canberra, ACT 2601, Australia
FEATURES             Location/Qualifiers
     source          1..4123
                     /organism="Oryza sativa"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:4530"
     gene            <1763..>3361
                     /gene="Osamt1-1"
     mRNA            <1763..>3361
                     /gene="Osamt1-1"
                     /product="ammonium transporter 1-1"
     CDS             1763..3361
                     /gene="Osamt1-1"
                     /codon_start=1
                     /product="ammonium transporter 1-1"
                     /protein_id="AAL05612.1"
                     /db_xref="GI:15705368"
                     /translation="MATCAADLAPLLGPVAANATDYLCNRFADTTSAVDATYLLFSAY
                     LVFAMQLGFAMLCAGSVRAKNTMNIMLTNVLDAAAGALFYLFGFAPAFGTPSNGFIG
                     KQFFGLKHMPQTGPDYDFFLFQWAFAIAAAGITSGSIAERTQFVAYLIYSAFLTGFVY
                     PVVSHWIWSADGWASASRTSGPLLFGSGVIDFAGSGVVHMVGGVAGLWGALIEGPRIG
                     RFDHAGRSVALKGHSASLVVLGTFLLMFGWYGFNPGSPTTILKTYGPAGGINGQWSGV
                     GRTAVTTLAGSVAALTTLFGKRLQTGHWNVVDVCNGLLGGFAAITAGCSVVDPWAAI
                     ICGFVSAWVLIGLNALAARLKFDDPLEAAQLHGGCGAWGILFTALFARQKYVEEIYGA
                     GRPYGLFMGGGGKLLAAHVIQILVIFGWVSCTMGPLFYGLKKLGLLRISABDETSGMD
                     LTRHGGFAYVVHDEDEHDKSGVGGFMLRSAQTRVEPAAAGCLQQQQPSVTNPERTTSQ
                     RRKKSRVSLPLRSRSSRHKFDPHI"
ORIGIN
```

```
Query Match   43.3%;  Score 738.2;  DB 8;  Length 4123;
Best Local Similarity   72.4%;
Matches  972;  Conservative  0;  Mismatches 368;  Pred. No. 7.2e-121;  Indels  3;  Gaps  1;

Qy  174 GCGCGCCGTGGACAGCAAGTCGTCGACAGCGCCTTCGCCGTCGACAACACCTACCTCCTC 233
        | | || |||||  |||||||| || ||     ||||||  | |||   |  |||||||||
Db 1823 GACTACCTGTGCAACCGGTTCGCCGACACGACGTCGGCGGTGGACGCGACGTACCTGCTC 1882

Qy  234 TTTTCCGCTACCTGCTGTTTTTCTATGCAGCCTGGGGTTCGCCATGCTCGCGCCGGCTCC 293
        || |||| ||||||||||| | | |||||||  ||| ||||||||||||||||||| | |
Db 1883 TTCTCGGCCTACCTGCTGTTCGCCATGCAGCTCGGGTTCGCCATGCTCGCGCCGGGTCG 1942
```

```
Qy  354 GGCCTCTTCTACTACCTCTTCGGCTTCGCTTCGCCTTCCCTCCAACGGCTTC 413
        | ||||||||||||||||||||||||||||| ||||||| |||||||||||||
Db  325 GCGCTCTTCTACTACCTCTTCGGCTTCGCCTTCGCCTTCGCCTCCAACGGCTTC 384

Qy  414 ATCGGTAAACATTCTTCGGCCTCAAGGACATCCCTCATCCTCTACGACTACAGCTAC 473
        ||||| || || ||||||||||||| ||||| ||  | |  ||| |||||| |||| |
Db  385 ATCGGGAAGCAGTTCTTCGGCCTCAAGCACATGCCGCAGACCGGGTTCGACTACGACTTC 444

Qy  474 TTCCTCTACCAATGGGCCTTCGCCATCGCCGCGGCATCACCAGCGGAAGCATCGCC 533
        ||||||| ||| |||||||||||||||||||| | ||||| | |||  ||||||||
Db  445 TTCCTCTTCCAGTGGGCCTTCGCCATCGCCGCGGGATCACGTCGGGCTCCATCGCC 504

Qy  534 GAACGCACACAGTTCGTGGCCTATCTCATCTACTCCGCCTTCCTCACCGGCTTCGTCTAT 593
        || || ||||||||||| ||||| ||||||||||||||||||||||||||| ||||||||
Db  505 GAGAGGACGCAGTTCGTCGCCTACCTCATCTACTCCGCCTTCCTCACCGGGTTCGTCTAC 564

Qy  594 CCGGTGGTGTCCCACTGGTTCTGGTCCCCAGACGGCTGGGCCTCTGCCTTTAAGATCACC 653
        |||||||||||||||||| |||||||| | |  ||| ||||||||||||     ||  ||
Db  565 CCGGTGGTGTCCCACTGGATCTGGTCCGCGGATGGGTGGGCCTCTGCCTCCCCGACGTCC 624

Qy  654 G--ACCGGCTATTTTCCACCGGGTAATAGACTTCGCCCGGTTCGGCGTAGTCCACATG 710
        |  ||| ||| ||| || ||| | | |||||||||| |  ||||||||||| ||||||||
Db  625 GGACCTTCGCTGTTCGGGTCCGGTGTGATCGACTTCGCCGGCTCCGGCGTCGTCCACATG 684

Qy  711 GTCGGGCCGAATAGCCGGCCTCATGGGAGCGCTGATCGAAGGCCCAAGAATGGGACGTTTC 770
        ||||| | | |   |||| || ||||| || | || |||||||||| |  |||| ||||||
Db  685 GTCGGCGGGGTGTCGCCGGGCTCTGGGGCGGCCTCATCGAGGGCCCGATCGGGAGGTTC 744

Qy  771 GATCATGCAGGACGAGCTGTGGCCTTGCTGGAGGCCACAGGCGTCCTTAGTAGTCCTGGGA 830
        || || || ||| | |  |||||| | |   ||||||||| ||||| | || || | |||
Db  745 GACCACGCCGGCGCATCGGTGGCGCTCAAGGGCCACAGCGCGTCGCTCGTCGTGCTTGGC 804

Qy  831 ACCTTCTTGCTTTGGTTCGGTTGGTACGGATTTAACCCCGGTTCATTTAACAAAATCCTA 890
        |||||| |||| ||| |||| |||   | || |||||||||| |  || |  | ||||| |
Db  805 ACCTTCCTGCTGCTGTTCGGTTCGGCGATTCAACCCCGGGTCGTTCACCACATCCTC 864

Qy  891 CTTACTTACGGTAACTCAGGAAATTACTACGGTCAATGGAGCGGCGTTGGCAGAACCGGCG 950
        |  || |||||  | |  || | | | | |||| ||||| |||| || ||| | ||||||
Db  865 AAGACGTACGGCCCGGCCGGCGGCATCAACGGGCAGTGGTCCGGAGTCGGCGCCGACGCC 924

Qy  951 GTCACCACTACCCTAGCCGGGGTCAACAGCTGCCTTGACACGCTATTCGGTAAACGGGTG 1010
        ||||||||  |||| ||||| ||||||| ||||||||||||| | |||||||  ||||| |
Db  925 GTGACGACGACCCTGGCCGGCAGCGTGGCGCGGCTCACCACGCTGTTCGGGAAGCGGGTC 984

Qy 1011 ATATCCGGTCACTGGAAGTCACCGATGTTCTGCAACGGCTGTGTTAGGCCGTTCGCGGCG 1070
        |  |  || |||||||| | ||| | | |||||||||| | || || ||||| ||  ||
Db  985 CAGACGGGGCACTGGAACGTGGTCGACGTCTGCAACGGCCTCCTGGGCGGGTTCGCCGCC 1044

Qy 1071 ATAACAGCCGGTTGCTCCGTGGTTGAGCCATGCAGCCATCGTATGCGGTTTGTTGTTGCT 1130
        || || ||||| || |||| ||||| || ||||| || | | | ||||  |  | |||||
Db 1045 ATCACCGCCGGGTGCAGCGTGCAGCCGTGCGTCGACCCGTGGGCCGCCATCATCTGCGGC 1104

Qy 1131 TCTATAGTATTAATAGCTTGCAACAAATTAGCAGAGAGGTTAAGTTCGACGATCCTCTG 1190
        || || |  ||  | || ||||   |||||| || || | ||||||||||||||| | ||
Db 1105 GCGTGGGTGCTCATCGGCCTCAACGCGCTGGCGGCCCGCCTCAAGTTCGACGACCCGCTC 1164

Qy 1191 GAGGGCGGCGCAGTTGACACGGTGGTGCACGGTGGGGGTGATATTCACGGCGGTGTTC 1250
        || | | || || | ||| ||||| || |||| |||||  ||  ||||| |  ||||||
Db 1165 GAGGCCGCCCAGCTCCACGGCGGGTGCGGCGCGTGGGGGATCCTCTTCACCGCGCTCTTC 1224

Qy 1251 GCAAAAAGGAGTATGTGAAGGATGTTTACGGGTTGGGGAGGCGCACGGGTTGCTCAATG 1310
        || |  || || || | ||||| | |||||| | || | || | ||||| || || ||
Db 1225 GCGAGGCAGAAGTACGTTGAGGAGATCTACGGCGCCGGCCGTCCGTACGGCGTCTTCATG 1284

Qy 1311 GGGGGTTGGTGGGAAGTTGCTGGCGGCACGTGATTCGATTCAGATTGCTTGGGGTGG 1370
        || || |||| || ||  ||| |||| ||| | || |||| ||||  | ||| ||||
Db 1285 GGCGGCGGCGGCAAGCTGCTCGCGGCCCACGTCATCCAGATCCTGGTCATCTTCGGGTGG 1344

Qy 1371 GTTAGTGCGACCATGGGACCCTTGTTTTGGGGGTTGAATAAACTGAGTGTTGAGGATT 1430
        || || ||||||||||||||| | || | ||| | |||| ||| || ||| |||| |
Db 1345 GTCAGCTGCACCATGGGACCGCTCTTCTACGGGCTCAAGAAGCTCGGCCTGCTGCGCATC 1404
```

```
Qy  3023  GTCAGCTGCACCATGGAGAGCTTGCGGGGACCTCTCTTCTACGGGGCTCAAGAAGCTCGGCCTGCTCCGCATC  3082
          ||  |||  |||||||||||  ||||  ||||  ||||||  |||  ||||  ||  |  ||  |||
Db

Qy  1431  TCTTCAGAGGATGAGAGCTTGCGGGGATGAACATGGACTCGCCATGAGGACGCTTGCTTATGCT  1490
          ||  ||||  ||||  |||  |||||||||  |||||||||||||  |||||||||||||
Db  3083  TCCGCCGAGGACGAGACGTCCGGCATGGACCTGACACGGCACGGCGGGTTCGCGTACGTC  3142

Qy  1491  TATGAGGATGATGAGACGCACAA  1513
          ||  ||||  ||  |||||
Db  3143  TACCACGACGAGGACGAGCACGA  3165
```

Search completed: March 9, 2004, 12:40:48
Job time : 6744 secs
```
Qy   294  GTCCGCGCCAAGAACACCATGAACATCATGCTCACCAACGTCCTGGACGCGTGCCGCCGGC       353
          |||| ||| ||||||||||||||| |||||||||||| |||||||| || ||  | |||||
Db  1943  GTGCGGGGCCAAGAACACGATGAACATCATGCTCACCAACGTGCTCGACGCGCGGGCCGGG      2002

Qy   354  GGCCTCTTCTACTACCTCTTCGGCTTCGGCCTTCGGCTTCCCCTCCAACGGCTTC          413
          ||||||||||||||||||||||||||||||||||||||||||  ||  | |||||
Db  2003  GCGCTCTTCTACTACCTCTTCGGCTTCGGCCTTCGGCCACGGCCGTCCAACGGCTTC      2062

Qy   414  ATCGGTAAACATTCTTCGGCCTCAAGGACATCCTTCATCCTCCTCACGACTACAGCTAC       473
          ||||||||  | |||||| ||||  |||  | ||||||| ||  || |||  |||||
Db  2063  ATCGGGAAGCAGTTCTTCGGCCTCAAGCACATGCCGCAGACCGGGTTCGACTACGACTTC      2122

Qy   474  TTCCTCTACCAATGGGCCTTCGCCATGCCGCCGCGGCCATCACCAGCGGAAGCATCGCC       533
          |||||||| | ||||||||||| |||||||| |||||  ||  || |||||||||||
Db  2123  TTCCTCTTCCAGTGGGCCTTCGCCATGGCCATGCCGCCGGAATCACGTCGGGCTCCATCGCC      2182

Qy   534  GAACGCACACAGTTCGTGGCCTATCTCATCTACTCCTCCTTCCTCACCGGCTTCGTCTAT       593
          |||||||| | ||||| ||||| |||  | |||| ||| |||||||||||||||||
Db  2183  GAGAGGACGCAGTTCGTGCGCTACCTCATCTACTCCTCCTTCCTCTCCTCACCGGTTCGTCTAC      2242

Qy   594  CCGGTGGTCTCCCACTGGTTCTGGTCCCAGACGGCTGGGCCTCTGCCTTTAAGATCACC       653
          |||||||||||||||||||||||||||||| || ||||||||| |  |||||||||||
Db  2243  CCGGTGGTGTCCCACTGGATCTGGTCGCGCGTGGGCCTCTGCCTCCCCGCAGTCC      2302

Qy   654  G--ACCGGCTATTTCCACCGGCTAATAGACTTCGCCGGTCCGGTGAGTCCACATG       710
          |   ||| ||| || ||||| ||||  || ||||| |||||||| |||||||
Db  2303  GGACCTCTGCTGTTCGGCTCCGGTGTCATCGACTTCGCCGGCGTCGTCACATG      2362

Qy   711  GTCGGCGGAATAGCCCGGCTATGGGGAGCCGTGATCGAAGGCCAAGAATGGACGTTC       770
          |||||||||| || |  || |||||| |||| ||||| |||| || | |||| |||
Db  2363  GTCGGCGGTGTCGCCGGGCTCGGGGCGCGCTCATCGAGGGCGCCCCGCATCGGGAGGTTC      2422

Qy   771  GATCATGCAGGACGAGCTGTGGCCTTGCGAGGCCACAGCGCGTCCTTAGTAGTCCTGGGA       830
          ||||||| || ||| |||||| ||  || |||| || || |||| || |||||||
Db  2423  GACCACGCCCGGCCGATCGGTCGCGCTCAAGGGCCCACAGCGCGTCGTCGCTCGTGCTTGGC      2482

Qy   831  ACCTTCCTTGGTTGGTTCGGTTGGTTGGTACGGATTTCATTTAACAAAATCCTA       890
          ||||||| || |  ||||| |  || ||||||| |||| || ||||||||
Db  2483  ACCTTCCTGCTGTTCGGCTGGTTGCGTTCAACCCCGGTCGTCACCACCATCCTC      2542

Qy   891  CTTACTTACCGTAACTCAGGAAATTACTACCGGTCAATGGAGCGCGGGTTGGCAGAACCGGG       950
          ||  ||  ||   || | ||||||  |||||| ||  ||  || || || |||
Db  2543  AAGACTACGGCCCCGGCCGGGGCATCAACGGGCCAGTGGTCCGGAGTCGGCCGCACCGCC      2602

Qy   951  GTCACCACTACCCTAGCGGGGTCAACAGCTTCCTGGTTGACCACGCTATTCGGTAAACGGGTG      1010
          || ||| ||  ||   | ||| |||||| |  |||| || || ||| || ||||| |
Db  2603  GTGACGACGACCCCTGCCGGCCGCAGCCTGCGCGCCTCACCACGCGTTGGCGGAAGCGGCTC      2662

Qy  1011  ATATCCGGTCACTGGAACGTGACCGATGTCTGCAACGGGCTGTGTAGGCGGTTCGGCGGCG      1070
          ||||||| |||| || |||  ||||| |||  ||| ||  |  ||||||  || ||||
Db  2663  CAGACGGGGCACTGGAACGTGGTCGACGTCTGCAACGGGGATCCTCTTCTCCACCGCGCTTC      2722

Qy  1071  ATAACAGCCGGTTGCTCCGTTGGTTGAGCCATGGGCAGCCATCGTATGCGGTTTTGTTGCT      1130
          ||  | | ||||  | | | |||||||||| |||  ||  ||  ||||| ||||||
Db  2723  ATCACCGCCCGGGTGCAGGTGCGTCGTCGACCCCGTGGGCGCGGATCATCCTGCGGGTTCGTCTCG      2782

Qy  1131  TCTATAGTATTAATAGCTTGAACAAATTAGCAGAGAAGGTTAAGTTCGACGATCCTCTG      1190
          |||  | || || ||| || || ||||| ||||||| ||| |||  |||||||||
Db  2783  GCGTGGGTGCTCATCGGCTTCAACGGCTTCAAGCGCCTCGCCGACGACCCGCTC      2842

Qy  1191  GAGGCGGCCGCAGTTGCACGGTGGTGTGGCACGTGGGGGTGATATTCACCGGCGTTGTTC      1250
          |||| ||  | |||| || || |||| ||||| || |||||| ||||  ||| ||||
Db  2843  GAGGCGCCCCAGCTCACGGCGGGGTTGCGGGCGGGGATCCTCTTCACCGCGCTTCTTC      2902

Qy  1251  GCAAAAAGGAGTATGTGAAGGAGTTTACGGGTTGGGAGGGGCACGGGGTTGCTCATG      1310
          || |  | ||| ||| || || |||||||||| |||||||||| | ||| |||||
Db  2903  GCGAGGCAGAAGTACGTGAAGGAGATCTTCGACGTGTCGCCGGTACGGCCGTGTTCATG      2962

Qy  1311  GGGGGTGGTGGGAAGTTGCTGGCGCGCACGTGATTCAGATTCGGTGATTGCTGGGGTGG      1370
          |||| |||| ||||| ||  |||||  || | ||| |||||| |||| || |||||
Db  2963  GGCGGCGCGGCCAAGCTGCTCGCCGCGCACGTCATCCTGGTGCTCGTGCTTCGGGGTGG      3022

Qy  1371  GTTAGTGCGACCATGGACCCCTTGTTTGGGGGTTGAATAAACTGAAGCTGTTGAGGATT      1430
```

Copyright   GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 9, 2004, 10:35:31 ; Search time 752 Seconds
                                         (without alignments)
                              9637.534 Million cell updates/sec

Title: US-10-033-109-3
Perfect score: 1706
Sequence: 1 gcacgagcactcccaccc........aaaaaaaaaaaaaaaaaa 1706

Scoring table: IDENTITY_NUC
             Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   N_Geneseq_29Jan04:*
             1: geneseqn1980s:*
             2: geneseqn1990s:*
             3: geneseqn2000s:*
             4: geneseqn2001as:*
             5: geneseqn2001bs:*
             6: geneseqn2002s:*
             7: geneseqn2003as:*
             8: geneseqn2003bs:*
             9: geneseqn2003cs:*
            10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1706	100.0	1706	7	AAL53985	Aal53985 cDNA enco
2	768.6	45.1	1991	7	AAL53986	Aal53986 cDNA enco
3	755.8	44.3	1506	3	AAC46516	Aac46516 Arabidops
4	755.8	44.3	1506	6	ABZ13398	Abz13398 Arabidops
5	755.8	44.3	1572	4	AAC44852	Aac44852 Arabidops
6	755.8	44.3	1748	2	AAQ88327	Aaq88327 Arabidops
7	705.6	41.4	1545	6	ABZ13685	Abz13685 Arabidops
8	700.4	41.1	1497	6	ABZ14082	Abz14082 Arabidops
9	221	13.0	1037	7	AAL53984	Aal53984 cDNA enco
c 10	121	7.1	445	6	ABL94034	Abl94034 Arabidops
c 11	89	5.2	985	6	ABQ44816	Abq44816 Oligonucl
12	89	5.2	985	6	ABQ44817	Abq44817 Oligonucl
13	87.6	5.1	1956	4	ABL11211	Abl11211 Drosophil
14	85.2	5.0	2805	4	AAD07130	Aad07130 Canine re
15	78.6	4.6	5296	4	ABL11210	Abl11210 Drosophil
16	77.8	4.6	799	2	AAV55831	Aav55831 Nucleotid
c 17	77.6	4.5	2803	4	AAD07132	Aad07132 Canine re
c 18	76.6	4.5	390	2	AAQ21833	Aaq21833 Randomisi
c 19	76.6	4.5	390	2	AAQ36859	Aaq36859 PCR prime
c 20	76.6	4.5	390	5	AAF76910	Aaf76910 Sequence
c 21	76.6	4.5	390	6	AAI72775	Aai72775 Oligo #7
22	76.4	4.5	1014	7	ADA70375	Ada70375 Rice gene
c 23	76	4.5	1050	4	AAD21685	Aad21685 Mutationa

c 24	76	4.5	1925	2	AAX90924	Aax90924 Epstein B
c 25	76	4.5	1926	3	AAA50254	Aaa50254 Epstein B
c 26	76	4.5	1926	3	AAF82902	Aaf82902 BBV tethe
c 27	76	4.5	2580	3	AAA75454	Aaa75454 Nucleotid
c 28	76	4.5	2580	6	AAI64275	Aai64275 Epstein-B
c 29	76	4.5	2849	4	AAD21684	Aad21684 Human ret
30	76	4.5	5452	2	AAX90923	Aax90923 Anti-sens
31	76	4.5	8705	2	AAZ23778	Aazz23778 Vector pS
c 32	76	4.5	9600	2	AAV21683	Aav21683 Vector pl
33	76	4.5	10285	6	ABS71027	Abs71027 pCEP-Xa-F
34	76	4.5	10285	6	ABS66453	Abs66453 Plasmid p
c 35	76	4.5	10380	2	AAZ22248	Aazz22248 Nucleotid
c 36	76	4.5	10596	2	AAQ51731	Aaq51731 Plasmid p
c 37	76	4.5	10596	2	AAX15650	Aax15650 Nucleotid
38	76	4.5	16080	2	AAT40348	Aat40348 Plasmid p
39	76	4.4	3849	4	AAA59553	Aaa59553 DNA clone
40	74.8	4.4	110000	4	AAF25795	Aaf25795 S. chryso
41	74.4	4.4	110000	4	AAI99682_27	Continuation (28 o
42	74.4	4.4	110000	4	AAI99682_28	Continuation (29 o
43	74.4	4.4	110000	4	AAI99683_27	Continuation (28 o
44	74.2	4.3	1248	3	AAZ87286	Aaz87286 S. venezu
45	74.2	4.3	1248	6	AAD39044	Aad39044 Streptomy

ALIGNMENTS

RESULT 1
AAL53985
ID   AAL53985 standard; cDNA; 1706 BP.
XX
AC   AAL53985;
XX
DT   18-FEB-2003 (first entry)
XX
DE   cDNA encoding a soybean high affinity ammonium transporter protein.
XX
KW   Herbicide; ammonium transporter protein; herbicide; transgenic plant;
KW   soybean; gene; ss.
XX
OS   Glycine max.
XX
FH   Key   Location/Qualifiers
FT   CDS   75..1577
FT   /tag= a
FT   /product= "Soybean high affinity ammonium transporter
FT   protein"
XX
PN   US2002142390-A1.
XX
PD   03-OCT-2002.
XX
PF   28-DEC-2001; 2001US-00033109.
XX
PR   28-AUG-1998; 98US-0098248P.
PR   27-AUG-1999; 99US-00384625.
XX
PA   (ALLE/) ALLEN S M.
PA   (RAFA/) RAFALSKI J A.
XX
PI   Allen SM, Rafalski JA;
XX
DR   WPI; 2003-102520/09.
DR   P-PSDB; AAO26533.
XX
PT   Novel ammonium transporter polypeptide useful for identifying enzymatic
PT   inhibitors, which is homologous to corn, soybean, wheat or rice ammonium
PT   transporter polypeptides.
XX
PS   Claim 3; Page 13-14; 27pp; English.
XX
CC   The invention relates to a novel ammonium transporter protein comprising
CC   90% homology based on the Clustal method compared to: a corn ammonium

```
CC   transporter polypeptide of 183 or 63 amino acids; a soybean ammonium
CC   transporter polypeptide of 500 or 486 amino acids; a wheat ammonium
CC   transporter polypeptide of 494 or 470 amino acids; and a rice ammonium
CC   transporter 497 amino acids fully defined in the specification. The
CC   isolated polynucleotide is useful for selecting an isolated
CC   polynucleotide that affects the level of expression of the ammonium
CC   transporter polypeptide in a plant cell. The ammonium transporter protein
CC   is useful for preparing antibodies which are useful for detecting the
CC   transporter protein in situ in cells or in vitro in cell extracts, and as
CC   targets to facilitate design and/or identify inhibitors of the enzymes
CC   that are useful as herbicides. The isolated polynucleotide is also useful
CC   for creating transgenic plants in which the polynucleotide is present at
CC   higher or lower levels than normal, and for designing and producing
CC   primer pairs which are useful in amplification or primer extension
CC   reactions. This polynucleotide sequence represents the cDNA encoding a
CC   soybean high affinity ammonium transporter protein of the invention
XX
SQ   Sequence 1706 BP; 359 A; 486 C; 442 G; 419 T; 0 U; 0 Other;

Query Match             100.0%;  Score 1706;  DB 7;  Length 1706;
Best Local Similarity   100.0%;  Pred. No.0;
Matches 1706;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GCACGAGCACTCCCAACCCCACCCGTAGTTCTACCACCTTCAGTCACGGCGTAATACA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GCACGAGCACTCCCAACCCCACCCGTAGTTCTACCACCTTCAGTCACGGCGTAATACA 60

Qy     61 CTAACCAACCACCATGTCGCTGCCTGTCTCCCGCGAACAACTGGCCCAACTTCTCG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CTAACCAACCACCATGTCGCTGCCTGTCTCCCGCGAACAACTGGCCCAACTTCTCG 120

Qy    121 GCCCAAACACCACAGAGCCTCGGCGCCGCCTCCCTTATCTGCGGCCATTCGCCCGCCG 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GCCCAAACACCACAGAGCCTCGGCGCCGCCTCCCTTATCTGCGGCCATTCGCCCGCCG 180

Qy    181 TGGACAGCAAGTTCGTCGACAACGGCCTTCGCCGTCGACAAACACCTTACCTCTCTTTCCG 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 TGGACAGCAAGTTCGTCGACAACGGCCTTCGCCGTCGACAAACACCTTACCTCTCTTTCCG 240

Qy    241 CCTACCTCGTTTTTCTATGCAGTCGGCTTCGGCTTCGGCTTCGCGCCGGCTCCGTCCGCGG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CCTACCTCGTTTTTCTATGCAGTCGGCTTCGGCTTCGGCTTCGCGCCGGCTCCGTCCGCGG 300

Qy    301 CCAAGAACACCATGAACATCATGCTCACCAAGTCCTGGACGCGTCCGCCGGCGGCCTCT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CCAAGAACACCATGAACATCATGCTCACCAAGTCCTGGACGCGTCCGCCGGCGGCCTCT 360

Qy    361 TCTACTACCTCTTCGGCCTTCGGCTTCGGCTCCCCTCCCCCTCCAACGGCTTCATCGGTA 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 TCTACTACCTCTTCGGCCTTCGGCTTCGGCTCCCCTCCCCCTCCAACGGCTTCATCGGTA 420

Qy    421 AACATTCTTCGGCCTCAAGGACATCCCTCCATCCTCCTACGACTACAGCTACTTCCTCT 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 AACATTCTTCGGCCTCAAGGACATCCCTCCATCCTCCTACGACTACAGCTACTTCCTCT 480

Qy    481 ACCAATGGGCCTTCGCCATGCCCATCGCGCCGGCATCACCAGCGGAAGCATCGCCGAACGCA 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 ACCAATGGGCCTTCGCCATGCCCATCGCGCCGGCATCACCAGCGGAAGCATCGCCGAACGCA 540

Qy    541 CACAGTTCGTGCGCCTATCTCATCTCATCCTCTCCCTCTCCTCACCGGCTTCGTCTATCCGGTGG 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 CACAGTTCGTGCGCCTATCTCATCTCATCCTCTCCCTCTCCTCACCGGCTTCGTCTATCCGGTGG 600

Qy    601 TCTCCCACTGGTTCTGGTTCCCAGACGGCTGGGCCTGCCTTTAAGATCACCGACCGGC 660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 TCTCCCACTGGTTCTGGTTCCCAGACGGCTGGGCCTGCCTTTAAGATCACCGACCGGC 660

Qy    661 TATTTTCCACCCGGCGTAATAGACTTCGGCGGTTCGGGCCGGTAGTCCACATGGTCGGCGGCGAA 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 TATTTTCCACCCGGCGTAATAGACTTCGGCGGTTCGGGCCGGTAGTCCACATGGTCGGCGGCGAA 720

Qy    721 TAGCCGGCCTATGGGGAGGCGCTGATCGAAGGCCCAAGAATGGACGTTTCGATCATGCAG 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 TAGCCGGCCTATGGGGAGGCGCTGATCGAAGGCCCAAGAATGGACGTTTCGATCATGCAG 780

Qy    781 GACGAGCTGTGGCCTTGCGGAGGCCACAGCCGTCCTTAGTAGTCCTGGGAACCTTCTTGC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 GACGAGCTGTGGCCTTGCGGAGGCCACAGCCGTCCTTAGTAGTCCTGGGAACCTTCTTGC 840

Qy    841 TTTGGTTCGGTTGGTACGGATTTAACCCCGGTCATTAACAAAATCCTACTTACTTACG 900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 TTTGGTTCGGTTGGTACGGATTTAACCCCGGTCATTAACAAAATCCTACTTACTTACG 900

Qy    901 GTAACTCAGGAAATTACTACGGTCAATGAGCGGTGGCAGAACCGCGGTCACCACTA 960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 GTAACTCAGGAAATTACTACGGTCAATGAGCGGTGGCAGAACCGCGGTCACCACTA 960

Qy    961 CCCTAGCCGGGGTCAACAGTCGCCTTGACCACGCTATTCGGTAAACGGGTGATATCCGGTC 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 CCCTAGCCGGGGTCAACAGTCGCCTTGACCACGCTATTCGGTAAACGGGTGATATCCGGTC 1020

Qy   1021 ACTGGAACGTGACCGATGCTGCAACGGGCTGTTAGGCGGTTCGCGGCGGATAACAGCCG 1080
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 ACTGGAACGTGACCGATGCTGCAACGGGCTGTTAGGCGGTTCGCGGCGGATAACAGCCG 1080

Qy   1081 GTTGCTCCGTGGTTGAGCCATGGGCAGCCATCGTATGCGGTTTGTTGCTTCTATAGTAT 1140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 GTTGCTCCGTGGTTGAGCCATGGGCAGCCATCGTATGCGGTTTGTTGCTTCTATAGTAT 1140

Qy   1141 TAATAGCTTGCAACAAATTAGCAGAGAAGGTTAAGTTCGACGATCCTCTGGAGGCGGCGC 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 TAATAGCTTGCAACAAATTAGCAGAGAAGGTTAAGTTCGACGATCCTCTGGAGGCGGCGC 1200

Qy   1201 AGTTGCACGGTGGGTGTGGCACGTGGGGGTGATATTCACCGGCGTTGTTCGGCAAAAAGG 1260
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 AGTTGCACGGTGGGTGTGGCACGTGGGGGTGATATTCACCGGCGTTGTTCGGCAAAAAGG 1260

Qy   1261 AGTATGTGAAGGAGGTTTACGGGGTTGGGGAGGGGCGCACGGGTTGCTCATGGGGGTGGTG 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 AGTATGTGAAGGAGGTTTACGGGGTTGGGGAGGGGCGCACGGGTTGCTCATGGGGGTGGTG 1320

Qy   1321 GGAAGTTGCTGGGCGGCCGCACGTGATTCAGATTCTGGTGATTGCTGGGTGGGTTAGTGCGA 1380
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 GGAAGTTGCTGGGCGGCCGCACGTGATTCAGATTCTGGTGATTGCTGGGTGGGTTAGTGCGA 1380

Qy   1381 CCATGGGACCCCTTGTTTGTTGGGGGTTGAATAAACTGTTGAAGCTGTTGAGGAGGTTCTTCAGAGG 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 CCATGGGACCCCTTGTTTGTTGGGGGTTGAATAAACTGTTGAAGCTGTTGAGGAGGTTCTTCAGAGG 1440

Qy   1441 ATGAGCTTGCGGGGATGGACATGACTCGCCATGGAGGCTTGCTATGCTTATGAGGATG 1500
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 ATGAGCTTGCGGGGATGGACATGACTCGCCATGGAGGCTTGCTATGCTTATGAGGATG 1500

Qy   1501 ATGAGACGGCACAAGCATGGATGCAGTTGAGGAGGTTGGGCCCAACGCGTCTTCCACAC 1560
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 ATGAGACGGCACAAGCATGGATGCAGTTGAGGAGGTTGGGCCCAACGCGTCTTCCACAC 1560

Qy   1561 CCACCACTGATGAATGAATGATCTTTTTTTCCCATATGCATGTCTCATTAGTAGTCAAACATTAAAT 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 CCACCACTGATGAATGAATGATCTTTTTTTCCCATATGCATGTCTCATTAGTAGTCAAACATTAAAT 1620

Qy   1621 TTGGATACATATTCCTTGTAAGGATTCAAACCTTGGTTACTTGTTACTTCTGTTAAAAAA 1680
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 TTGGATACATATTCCTTGTAAGGATTCAAACCTTGGTTACTTGTTACTTCTGTTAAAAAA 1680

Qy   1681 AAAAAAAAAAAAAAAAAAAAAAAAAA 1706
          ||||||||||||||||||||||||||
Db   1681 AAAAAAAAAAAAAAAAAAAAAAAAAA 1706

RESULT 2
AAL53986
ID   AAL53986  standard; cDNA; 1991 BP.
XX
AC   AAL53986;
XX
```

```
DT   18-FEB-2003   (first entry)
XX
DE   cDNA encoding a wheat high affinity ammonium transporter protein.
XX
KW   Herbicide; ammonium transporter protein; herbicide; transgenic plant;
KW   wheat; gene; ss.
XX
OS   Triticum aestivum.
XX
FH   Key            Location/Qualifiers
FT   CDS            71..1555
FT                  /*tag= a
FT                  /product= "Wheat high affinity ammonium transporter
FT                  protein"
XX
PN   US2002142390-A1.
XX
PD   03-OCT-2002.
XX
PF   28-DEC-2001; 2001US-00033109.
XX
PR   28-AUG-1998;  98US-0098248P.
PR   27-AUG-1999;  99US-00384625.
XX
PA   (ALLE/) ALLEN S M.
PA   (RAFA/) RAFALSKI J A.
XX
PI   Allen SM, Rafalski JA;
XX
DR   WPI; 2003-102520/09.
DR   P-PSDB; AAO26534.
XX
PT   Novel ammonium transporter polypeptide useful for identifying enzymatic
PT   inhibitors, which is homologous to corn, soybean, wheat or rice ammonium
PT   transporter polypeptides.
XX
PS   Claim 3; Page 16-17; 27pp; English.
XX
CC   The invention relates to a novel ammonium transporter protein comprising
CC   90% homology based on the Clustal method compared to: a corn ammonium
CC   transporter polypeptide of 183 or 63 amino acids; a soybean ammonium
CC   transporter polypeptide of 500 or 486 amino acids; a wheat ammonium
CC   transporter polypeptide of 494 or 470 amino acids; and a rice ammonium
CC   transporter 497 amino acids fully defined in the specification. The
CC   isolated polynucleotide is useful for selecting an isolated
CC   polynucleotide that affects the level of expression of the ammonium
CC   transporter polypeptide in a plant cell. The ammonium transporter protein
CC   is useful for preparing antibodies which are useful for detecting the
CC   transporter protein in situ in cells or in vitro in cell extracts, and as
CC   targets to facilitate design and/or identify inhibitors of the enzymes
CC   that are useful as herbicides. The isolated polynucleotide is also useful
CC   for creating transgenic plants in which the polynucleotide is present at
CC   higher or lower levels than normal, and for designing and producing
CC   primer pairs which are useful in amplification or primer extension
CC   reactions. This polynucleotide sequence represents the cDNA encoding a
CC   wheat high affinity ammonium transporter protein of the invention
XX
SQ   Sequence 1991 BP; 310 A; 633 C; 624 G; 424 T; 0 U; 0 Other;
```

```
Query Match      45.1%;  Score 768.6;  DB 7;  Length 1991;
Best Local Similarity   73.8%;  Pred. No. 1.6e-148;
Matches  991;  Conservative   0;  Mismatches 349;  Indels   3;  Gaps   1;

Qy  174 GCGCGCCGTGGACGCAGCAAGTTCGTCGACACGGCCTTCGCCGTCGACAACACCTACCTCTC 233
Db  134 GACTACCTGTGCAACAGAGTTCGCCGACACCACGTCCGCCGGTGACTCCACCTACCTGCTC 193

Qy  234 TTTTCCGCCTACCTCGTTTTTTCTATGCAGCTCGGCTTCGCCATGCTTCGCGCGCCTCC 293
Db  194 TTCTCGGCCTACCTCGTTCGCCTTCGCCCATGCTCGGCTTCGCCATGCTCGCGCGGCTCC 253

Qy  294 GTCCCGGCCAAGAACAACACCATGAACATCATGCTCACCAACGTCCTGGACGCTCCTGA 353
```

```
Db  254 GTCCGGGCCAAGAACAACACCATGAACATCATGCTCACCAACGTGCTGCAGCCGCCGCCGGC 313
Qy  354 GGCCTCTTCTACTACCTTCTTCGGCTTCGGCTTCGGCTTCCCCCTCCAACGGCTTC 413
Db  314 GCGCTCTTCTACTACCTTCTTCGGCTTCGGCTTCGGCTTCCCTTCGGGACGCCGTGAACGGCTTC 373
Qy  414 ATCGGTAAACATTTCTTCGGCTCAAGGACATCCCTTCATCCTCTACGACTACAGCTAC 473
Db  374 ATCGGGAAGCACTTCTTCGGCCTCAAGGACATGCCGCAGGACCCGGCTTCGACTACAGCTTC 433
Qy  474 TTCCTCTACCAATGGCCTTCGCCATGCTCGCCGCCGCCGGCATCACCAGCGGAAGCATCGCC 533
Db  434 TTCCTCTTCCAGTGGGCCTTCGCCATGGCCCATCGCCGCCGCCGGCATCACCTCACCTCCCGGCTCCATCGCC 493
Qy  534 GAACGGCACACAGTTCGTGGCCTATCTCATCTACTCCCTTCCTCACCGGCTTCGTCTCTAT 593
Db  494 GAGAGGACGCAGTTCGTCGCCGTATCTCATCTACTCCGGCCTTCCTTACGGGATTCGTCTAC 553
Qy  594 CCGGTGGTCTCCCACTGGTTCTGGCTCGGGCTCTGGGCCTTCTGCCTTTAAGATCACC 653
Db  554 CCGGTCGTGTCCCACTGGATCTGGTTCCGTGACGGGCTTGGGCCTCCCGCCGGCCCGCCACGTCC 613
Qy  654 GACCGGCTATTTTCCA---CCGGCGTAATAGACTTCGCCGGTTCGGCGTAGTCCACATG 710
Db  614 GGCCCGCCTGCTCTTCAAGTCCGGCGTGATCGACTTCGCCGGTTCGGCGTCGTGCACATG 673
Qy  711 GTCGGCGGAATAGCCGGCCTATGGGGAGGCGTGATCGAAGGGCCCAAGAATGGGACGTTC 770
Db  674 GTCGGCGGCATCGCCGGCCTTCTGGGGGCGTCATCGAGGGGCCCCGCCATCGGCCGGTTC 733
Qy  771 GATCATGCAGGACGAGCTGTGGCCTTGCGAGGCCACAGCGCGTCCTAGTAGTCCTGGGA 830
Db  734 GACCACGCCGGGCCGCTCGGTGGTGGCGCTCAAGGGCCAAGCGCGTCGCTCGTCGTCGTGGGG 793
Qy  831 ACCTTCTTGTTTGGTTCGGTTCGGCTGGTACGGGTTCAATTTAACAAAATCCTA 890
Db  794 ACCTTCCTGCTCTGGTCTCGGTTCGGCTGGTACGGGTTCAACCCGGGGTTCCTTCGTCGTCACCATCCTC 853
Qy  891 CTTACTTACGGTAACTCAGAAATTACTACGGTCAATGGAGCGCGGTTGGCAGAACCGCG 950
Db  854 AAGTCGTACGGCCCCGCCGGGAGCATCAACGGGCAGTGGTCGGGCGGTGGCCGCACCGGCC 913
Qy  951 GTCACCACTACCCTAGCGGGGTCAACAGCTGCCTTGACCACGCTATTCGGTAAACGGGGTC 1010
Db  914 GTGACGACGACGCGCGTGGCGCAGCGTGGCGGGCGCTCACGACCGGTTCGGAAGCGGCGTC 973
Qy  1011 ATATCCGGTCACTGGAACGTGACCGATGTCGCAACGGCGTGTTAGGCGGTTCGGCGGCG 1070
Db  974 CAGACGGGGCACTGGAACGTGGTGGACGTCTGCAACGGCCTGCTCGGCGGGGTTCGCGGCC 1033
Qy  1071 ATAACAGCCGGTTGCTCCGGTTGAGCCATGGCAGCCATCGTATGCGGTTTGTTGCT 1130
Db  1034 ATCACCGCCGGGGTGCAGCGGTGGTGCGACCCGTGGGGCCGTCATCTGCGGCTTCGGTCTCC 1093
Qy  1131 TCTATAGTATTAATAGCTTGCAACAAATTAGCAGAGAAGGTTAAGTTCGACGATCCTCTG 1190
Db  1094 GCCTGGGTGCTCATCGGGCTCAACGCGCTCGCCGGCGCGCTCGCCTCCAAGTTCGAGGGCTG 1153
Qy  1191 GAGGGCGGCGCAGTTGCACGGTGGTGCACGTGGGGGGTGATATTCACGGCGTTGTTC 1250
Db  1154 GAGGCGGGCAGCTGCAGTCGCACGGCGCGGCTGCGGGCGGTGGGGGATCATCTTCACCGGCGCGTTC 1213
Qy  1251 GCAAAAAAGGAGTATGTGAAGGAGGTTTACGGGTTCGGGAGGGCGCACGGTTGCTCATG 1310
Db  1214 GCCAAGAGCAGTACGTGGAGGAGATCTACGGCGGCGGCTACGGGCTGCTGCTGTTCCTG 1273
Qy  1311 GGGGGTGGTGGGAAGTTGCTGCGGCGCACGTGATTCAGATTCTGGTGATTGCGGTGGTGG 1370
Db  1274 GGCGGCGGCGGGCGCGGCGTGCTGCGGCGGCCACCATCGTGCAGATCCTCGTCATCGCCGGCTTC 1333
Qy  1371 GTTAGTGCGACCATGGGACCCTTGTTTTCGGGGGTTGAATAAACTGAAGCTGTTGAGGATT 1430
Db  1334 GTGAGCTGCACCATGGGGCCGCCTCTTCTTGGCGCTCAAGAGAAGCTGGGCGTCCCGCATC 1393
```

```
Qy    1431  TCTTCAGAGGATGAGCTTGCGGGGATGGACATGACTCGCCATGAGAGGCTTTGCTTATGCT 1490
                  ||||||  ||||| | ||||||||||||||| ||||| ||||||||  |||| |||
Db    1394  TCGGCCGGAGGACGAGATGGCCGGCCATGAGGACCTGACCCGGCACGGGTTCGCCTACGTC 1453

Qy    1491  TATGAGGATGATGAGACGGCCACAA 1513
                  ||  || |||  |||||  ||||
Db    1454  TACCACGACGACGACGAGCACGA 1476
```

RESULT 3
AAC46516
ID   AAC46516 standard; DNA; 1506 BP.
XX
AC   AAC46516;
XX
DT   18-OCT-2000   (first entry)
XX
DE   Arabidopsis thaliana DNA fragment SEQ ID NO: 50425.
XX
KW   Hybridisation assay; genetic mapping; gene expression control;
KW   protein identification; signal transduction pathway; metabolic pathway;
KW   promoter; termination sequence; ss.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-00301439.
XX

PR	25-FEB-1999;	99US-0121825P;	PR	10-JUN-1999;	99US-0138847P.
PR	05-MAR-1999;	99US-0123180P;	PR	14-JUN-1999;	99US-0139119P;
PR	09-MAR-1999;	99US-0123548P;	PR	16-JUN-1999;	99US-0139452P;
PR	23-MAR-1999;	99US-0125788P;	PR	17-JUN-1999;	99US-0139453P;
PR	25-MAR-1999;	99US-0126264P;	PR	18-JUN-1999;	99US-0139492P;
PR	29-MAR-1999;	99US-0126785P;	PR	18-JUN-1999;	99US-0139454P;
PR	01-APR-1999;	99US-0127462P;	PR	18-JUN-1999;	99US-0139455P;
PR	06-APR-1999;	99US-0128234P;	PR	18-JUN-1999;	99US-0139456P;
PR	08-APR-1999;	99US-0128714P;	PR	18-JUN-1999;	99US-0139457P;
PR	16-APR-1999;	99US-0129845P;	PR	18-JUN-1999;	99US-0139458P;
PR	19-APR-1999;	99US-0130077P;	PR	18-JUN-1999;	99US-0139459P;
PR	21-APR-1999;	99US-0130449P;	PR	18-JUN-1999;	99US-0139460P;
PR	23-APR-1999;	99US-0130510P;	PR	18-JUN-1999;	99US-0139461P;
PR	23-APR-1999;	99US-0130891P;	PR	18-JUN-1999;	99US-0139462P;
PR	28-APR-1999;	99US-0131449P;	PR	18-JUN-1999;	99US-0139463P;
PR	30-APR-1999;	99US-0132048P;	PR	18-JUN-1999;	99US-0139750P;
PR	30-APR-1999;	99US-0132407P;	PR	21-JUN-1999;	99US-0139763P;
PR	04-MAY-1999;	99US-0132484P;	PR	23-JUN-1999;	99US-0139817P;
PR	05-MAY-1999;	99US-0132485P;	PR	23-JUN-1999;	99US-0139899P;
PR	06-MAY-1999;	99US-0132486P;	PR	24-JUN-1999;	99US-0140353P.
PR	06-MAY-1999;	99US-0132487P;	PR	28-JUN-1999;	99US-0140354P.
PR	07-MAY-1999;	99US-0132863P;	PR	29-JUN-1999;	99US-0140823P;
PR	11-MAY-1999;	99US-0134941P;	PR	30-JUN-1999;	99US-0140991P;
PR	14-MAY-1999;	99US-0134256P;	PR	01-JUL-1999;	99US-0141287P;
PR	14-MAY-1999;	99US-0134218P;	PR	01-JUL-1999;	99US-0141842P;
PR	14-MAY-1999;	99US-0134219P;	PR	01-JUL-1999;	99US-0142154P;
PR	14-MAY-1999;	99US-0134221P;	PR	02-JUL-1999;	99US-0142055P;
PR	14-MAY-1999;	99US-0143370P;	PR	06-JUL-1999;	99US-0142390P;
PR	18-MAY-1999;	99US-0134768P;	PR	08-JUL-1999;	99US-0142803P;
PR	19-MAY-1999;	99US-0134941P;	PR	09-JUL-1999;	99US-0142920P;
PR	20-MAY-1999;	99US-0135124P;	PR	12-JUL-1999;	99US-0142977P;
PR	21-MAY-1999;	99US-0135353P;	PR	13-JUL-1999;	99US-0143542P;
PR	24-MAY-1999;	99US-0135629P;	PR	14-JUL-1999;	99US-0143624P;
PR	25-MAY-1999;	99US-0136021P;	PR	15-JUL-1999;	99US-0144005P;
PR	27-MAY-1999;	99US-0136392P;	PR	16-JUL-1999;	99US-0144085P;
PR	28-MAY-1999;	99US-0136782P;	PR	16-JUL-1999;	99US-0144086P;
PR	01-JUN-1999;	99US-0137222P;	PR	19-JUL-1999;	99US-0144325P;
PR	03-JUN-1999;	99US-0137528P;	PR	19-JUL-1999;	99US-0144331P;
PR	04-JUN-1999;	99US-0137502P;	PR	19-JUL-1999;	99US-0144332P;
PR	07-JUN-1999;	99US-0137724P;	PR	19-JUL-1999;	99US-0144333P;
PR	08-JUN-1999;	99US-0138094P;	PR	19-JUL-1999;	99US-0144334P;
PR	10-JUN-1999;	99US-0138540P;	PR	19-JUL-1999;	99US-0144335P;
			PR	20-JUL-1999;	99US-0144352P;
			PR	20-JUL-1999;	99US-0144632P;
			PR	21-JUL-1999;	99US-0144884P;
			PR	21-JUL-1999;	99US-0144814P;
			PR	22-JUL-1999;	99US-0145086P;
			PR	22-JUL-1999;	99US-0150088P;
			PR	22-JUL-1999;	99US-0145085P;
			PR	22-JUL-1999;	99US-0145087P;
			PR	22-JUL-1999;	99US-0145089P;
			PR	23-JUL-1999;	99US-0145192P;
			PR	23-JUL-1999;	99US-0145145P;
			PR	23-JUL-1999;	99US-0145218P;
			PR	26-JUL-1999;	99US-0145224P;
			PR	27-JUL-1999;	99US-0145276P;
			PR	27-JUL-1999;	99US-0145913P;
			PR	27-JUL-1999;	99US-0145918P;
			PR	28-JUL-1999;	99US-0145919P;
			PR	02-AUG-1999;	99US-0145951P;
			PR	02-AUG-1999;	99US-0146386P;
			PR	02-AUG-1999;	99US-0146388P;
			PR	03-AUG-1999;	99US-0146389P;
			PR	04-AUG-1999;	99US-0147038P;
			PR	04-AUG-1999;	99US-0147204P;
			PR	05-AUG-1999;	99US-0147302P;
			PR	05-AUG-1999;	99US-0147192P;
			PR	06-AUG-1999;	99US-0147260P;
			PR	06-AUG-1999;	99US-0147303P;
			PR	09-AUG-1999;	99US-0147416P;
			PR	09-AUG-1999;	99US-0147493P;
			PR	09-AUG-1999;	99US-0147935P;
```

Db 44 ATGCCACGGCGCGGCCAACTACATCTGTGGCCAGTTAGGCGACGTCAACAACAAATTTA 103

Qy 196 TCGACACGGCCTTCGCCGTCGACAACACTACCTCCTTCTCCTCTCCGCCTACCTCGTTTTT 255
Db 104 TCGACACCGCCTTCGCTATAGACAACAACTTACCTTCTCTCCGCCTACCTCGTCTTCT 163

Qy 256 CTATGCGAGCTCGGCCTTCGCGGCGCTCCGTCGCTCCGCCGGAGAACACCATGA 315
Db 164 CTATGCAGCTTGGCTTCGCTATGCTCTGTGCCGGTTCCGTGAGGCAAGAATACTATGA 223

Qy 316 ACATCATGCTCACCAACGTCCTGACGGCGTCCGCGGCGGCCTCTTCTACTACCTCTTCG 375
Db 224 ACATCATGCTTACCAACGTCCTTGAGGCGCTGCAGCCGGTGCTCTTCTATTATCTGTTTG 283

Qy 376 GCTTCGCCTTCGCTTTCGGCTTCCCCCTCCAACGGCTTCATCGTGAAACATTCTTCGGCC 435
Db 284 GCTACGCCTTTGCCTTTGGATCTCCGTCCAATGGTTCATCGGTAAACACTACTTGGTC 343

Qy 436 TCAAGGACATCCCTTCATCCTCCTACGACTACAGCTACTTCCTCTACCAATGGGCCTTCG 495
Db 344 TCAAAGACATCCCCACGGCCTCTGACTACTCCAACTTTCTACTACCAATGGGCCTTTG 403

Qy 496 CCATCGCCGCGCCGCATCACCAGCGGGAAGCATCGGCCGAACGCACACAGTTCGTGGCCT 555
Db 404 CAATCGCTGCGGGCTGGAATCACAAGTGGCTCGATCGCTGAACCGGACACAGTTCGTGGGCTT 463

Qy 556 ATCTCATCTACTCCTTCCTTCACCGGCTTCGTCTATCCGGTGGTCTCCCACTGGTTCT 615
Db 464 ACCTAATCTATTCCTTCCTTTCTTCTTAACCGGTTTGTTTTTACCCGGTCGTCTCTCACTGGTTCT 523

Qy 616 GGTCCCCAGACGGCTGGGCCTTCTGCTT---TAAGATCACCGACCGGCTATTTTCCACCG 672
Db 524 GGTCAGTTGATGGATGGGCCAGCCCGTTCCGTACCGATGGAGATTTGCTTTTCAGCACCG 583

Qy 673 GCGTAATAGATCTTCCCGGTTCCGGCGTAGTCCACATGGTCGCGGAATAGCCGGCCTAT 732
Db 584 GAGCGGATAGATTTCGCTGGGGTCCGGTGTTGTCATATGGTCGGAGGTATCGGCTGGGACTCT 643

Qy 733 GGGGAGCGGCTGATCGAAGGCCCAAGAAATGGGACGTTCGATCATGCAGGACGAGCTGTGG 792
Db 644 GGGGTGCGGCTCATCGAAGGTCCACGACTTGGCCGGTTCGATAACGGAGGCCGTGCCATCG 703

Qy 793 CCTTGCGAGGCCACAGCGCGGCGGTCCTTAGTAGTCCTGGGGAACCTTCTTGCTTGGTTCGGTT 852
Db 704 CTCTTCGTGGCCACTCGGCGTCACTTGTTGTCCTTGGAACATTCCTCGTCGTTTGGAT 763

Qy 853 GGTACGGATTTAACCCCGGTTCATTGCATTAACAAAATCTACTTACTTACGGTAACTCAGGAA 912
Db 764 GGTACGGATTTAACCCGGTTCCTTCAACAGATCCTAGTCACGTACGAGA---CAGGCA 820

Qy 913 ATTACTACGGTCAATGGAGCGCGGTTGGCAGAACCGCGGTCACCACTACCCTAGCCGGGGGT 972
Db 821 CATACAACGGCCAGTGAGCGGCGGTCGGACGGACAGCGTGTCACAACAAGTTAGCTGGCT 880

Qy 973 CAACAGCTGCCTTGCCTTGACCACGCTATTCGGTAAACGGGTGATATCCGGTCACTGGAACGTGA 1032
Db 881 GCACCGGCGCGCTGACAACCCCATTTGGAAAGTCTACTCTCGGGACATTGGAACGTCA 940

Qy 1033 CCGATGTCGCAACGGGCTGTGTTAGGCGGTTCGCGGATAACAGCCGGTTGCTCCGTGG 1092
Db 941 CTGATGTGATGCAACGGCGCCTTCCTCCTGGAGGGTTTGCAGCCATAATCGTGGCGCTCTGTCG 1000

Qy 1093 TTGAGGCCATGGCAGCCATCGTATGCGGTTTGTTGCTTCTATAGTATTAATAGCTTGCA 1152
Db 1001 TTTGAGGCCATGGCAGCCGATCATCGTGCGGGTTCGGTGGCGGCCCTAGTTCCTCCTGGATGCA 1060

Qy 1153 ACAAATTAGCAGAGAAGGTTAAGTTCGACGATCCTCTGGAGGGCGGCCAGTTGCACGGTG 1212
Db 1061 ACAAGCTCGCTGAGAAGTCAAATACGACGACCCTCTTGAGGCAGCACAACTACACGGTG 1120

Qy 1213 GGTGCGACGTGGGGGTGATATTCACGGCGGTTGTTCGCCAAAAAGGAGGATGTGAAGG 1272
Db 1121 GTTGCGGTGCGGTGGGGACTAATATTCACGGCTCTTCCGCTCAAGAAAAAGTACTTGAACC 1180

Query Match 44.3%; Score 755.8; DB 3; Length 1506;
Best Local Similarity 71.7%; Pred. No. 6.5e-146;
Matches 1036; Conservative 0; Mismatches 397; Indels 12; Gaps 3;

Qy 136 ACGCCCTCCGCCGCGCCTCCTTATCTGCGGCCATTCGCCGCCGCCGTGACAGCAAGTTCG 195
            |||| ||   ||| | |||    |||||||| |||||| ||  ||| ||| |||  |||

Db 136 ACGCCCTCCGCCGCGCCTCCTTATCTGCGGCCATTCGCCGCCGCCGTGACAGCAAGTTCG 195

```
CC   information supplied to Derwent by the European Patent Office
XX
SQ   Sequence 1506 BP; 307 A; 404 C; 388 G; 407 T; 0 U; 0 Other;

     Query Match          44.3%;    Score 755.8;  DB 6;   Length 1506;
     Best Local Similarity 71.7%;   Pred. No. 6.5e-146;
     Matches 1036; Conservative 0; Mismatches 397; Indels 12; Gaps 3;

Qy   136 ACGGCTCCGCCGCCGCCTCCTTATCTGCGGCCATTCGCCGCCGTGCCGGCCGGACAGCAAGTTCG 195
          ||||| ||||| || ||| |||||   ||||||||| || |||||| ||  ||||| || || |||
Db    44 .ATGCCACGGCGGCGGCCCAACTACATCTGTGGCCAGTGCGACGTCAACAACAAAATTA      103

Qy   196 TCGACACGGCCTTCGCCGTGCAGACACCTACCTCCTCTTTCCGCCTACCTCGTTTTT       255
          ||||||| || ||| ||  |||| ||| |||||| || || ||||| ||||||| | ||
Db   104 TCGACACCGGCTTTCGCTATAGACAACACTTACCTTCTCTCCGGCTACCTTGTCTTCT      163

Qy   256 CTATGCAGCTCGGCTTCGCCATGCTCGGCGCCTCCGTCCGGCCGCCAAGAACACCATGA     315
          ||||||||||| ||||| || ||||| |||  | ||||  || || ||||||| ||||||
Db   164 CTATGCAGCTTGGCTTCGCTGTGCGGGTTCCGTGCGGTTCCTGAGAGCCAAGATAACTATGA   223

Qy   316 ACATCATGCTCACCAAGTCCTGGAGCGTGCCGCCGGCGGCCTCTTCTACTACCTCTTCG     375
          ||||||||||| |||||  ||| |||  | |||| ||| || |||||| || || | | ||
Db   224 ACATCATGCTTACCAACGTCCTTGACGCTCCAGCCGCGGTGCTCTCTTCTATTATCTGTTTG   283

Qy   376 GCTTCGCCTGCTTTCGGCTCCCCTCCGTCCCCCTCCAAGGCTTCATCGGTAAACATTCTTCCGGCC  435
          |||| ||  |||| |  |||| |  ||| |    |||| |||| ||||| ||||||| ||| | |||
Db   284 GCTACGGCCTTTGCCTTTGGATCTCCGTCCAATGGTTTCATCGGTTCATCGGTAAACACTACTTTGGTC  343

Qy   436 TCAAGGACATCCTTCATCCTCCTACGACTACAGCTACTTCCTACCAATGGGCCTTCG       495
          ||||| |||| | ||  || || || ||| | ||||| |||| || |||| |||  ||
Db   344 TCAAAGACATCCCCACAGGCCTCTGCTGACTACTCCAACTTTCTCACCAATGGGCCTTTG      403

Qy   496 CCATCGCCGCCGCCGGCATCACCAGCGGAAGCATCGCCACAGTTCGTCGTGGCCT        555
          |||| | ||||| || |||||| ||||||| |  |  ||||||| | ||||| |||
Db   404 CAATCGCTGCGGCTGGAATCACAAGTGGCTCGATCGTCGACGGACACAGTTCGTGGCTT     463

Qy   556 ATCTCATCTCACTCCTCCCTCCGTCGTCTATCCGGTCGTCGTCTCCCACTGGTTCT        615
          |||| | |||| || | || || |||| |   |||  ||||||  || |  | |||||
Db   464 ACCTAATCTATTCCTCTTCTTACCGGGTTTGTTGTTTTACCCGGTCGTCTCTCCACTGGTTCT    523

Qy   616 GGTCCCCAGACGGCTGGGCCTCTGCCTT---TAAGATCACCGACCGGCTATTTTCCACCG      672
          ||   ||| ||||| |||  |  | ||||   |||| ||| || || ||||||   ||  |
Db   524 GGTCAGTTGATGGATGGGCCAGCCGTTCCGTCGGTACCGGACCGGATTGCTTTTCAGCACCG      583

Qy   673 GGTAATAGACTTCGCCGGTTCGGCGTAGTCCACATGGTCGGCGGAATAGCCGGCCTAT     732
          ||| | || || || || ||| || |  | ||  |||| |||| | ||||| || || |
Db   584 GAGCGATAGATTTCGCTGGGTTCGGTGTGTTCATATGGTCGGAGGTATCGTCGGACTCT      643

Qy   733 GGGGAGCGCTGATCGAAGGCCCAAGAATGGGACGTTCGATCATGCAGGACGAGCGTGTGG     792
          |||| |||||| || ||| ||| || | ||| | ||||| || || |||| |||||| ||
Db   644 GGGGTGCGCTCATCGAAGGTCCACGACTTGGCCGGTTCGATAACCGGAGGGCCGTGCCATCG     703

Qy   793 CCTTGGGAGGCCACAGCGCGCGTCCTTAGTAGTCCTGGGAAACTTCTTGCTTTGGTTCGGTT    852
          | |  | |||| |||  |||  | |||  |||  | || ||| | ||||||| |||| ||||||
Db   704 CTCTTCGTGGCGCACTCGGGCGTCACTTGTTGTCCTTGGAACCATTCCTCCTCGGTTTGGAT   763

Qy   853 GGTACGGATTTAACCCCGGTTCATTTAACAAATCCTACTTACGGTAACTCAGGAA        912
          |||||||||| | ||  || ||  |||| |||||| || |||| ||| |||  |||
Db   764 GGTACGGATTTAACCCCGGTTCCTTCAACAAGATCCTAGT-CACGTACGAGA---CAGGCA  820

Qy   913 ATTACTACGGTCAATGAGCGCGGTTGGCAGAACGCGGTCAGCACCACCTACCCTAGCGGGGT     972
          || | | ||| | || ||||| ||||||||||| || ||||| || ||| || |||||  |
Db   821 CATCACACGGCCAGTGGAGCGCGGTCGGACGGACGAGCGTCACACAACAACGTTAGCTGGCT    880

Qy   973 CAACAGCTGCCTTGACCACGCTATTCGGTAAACGGGTGATATCCGGTCACTGGAACGTGA    1032
          ||||| | ||||| || |||  || || | ||| ||||| | |  || |||| ||| | |
Db   881 GCACCGGGGCGCTGACAACCCTATTTGGAAACGTCTACTCTCGGGACATTGGAACGTCA     940

Qy  1033 CCGATGTCTGCAAACGGGTGTAGGCGGTTTCGCGGGAATAACAGCGGTTGCTCCGTGG    1092
          | ||||  |||||| || ||  || |||  | |  ||| || ||||| |||||  |||
Db   941 CTGATGTATGCAACGACCGCCTCCTGGAGGGTTTGCAGCCATAACTGTGGCTGCTCTGTCG   1000

Qy  1093 TTGAGCCATGGGCAGCCATCGTATGGCGTATGGGTTTGTTGCTTCATAGTATTAATAGCTTGCA   1152
          |||| |||| || ||| || | ||||| |||  ||| |||| ||||| || || || |||||
```

```
Db  1001 TTGAGCCATGGGCTGCGATCATCTGCGGGGTTCGTGGCGGCCCCTAGTCCTTCCTCGGATGCA 1060
         ||||||||||||| ||||| ||||| ||| |||| |||     ||    ||  | ||||||||
Qy  1153 ACAAATTAGCAGAGAAGGTTAAGTTCGACGATCCTCTGGAGGCGGCCGCAGTTGCACGGTG 1212
         |||||    |||||  ||||  ||||| ||||| | ||||||||| ||| ||| | ||||||
Db  1061 ACAAGCTCGCTGAGAAGCTCAAATACGACGACCCTCCTTGAGGCAGCACAACTACACGGTG 1120

Qy  1213 GGTGTGGCACGTGGGGCGGTGATATTCACGGGCGTTGTTCGCAAAAAGGAGTATGTGAAGG 1272
         ||||| |||||| |||| |||||||||||| ||  | || |||||| || || || ||  |
Db  1121 GTTGCGGTGCGTGGGGACTAATATTCACGGCGTCCTCTTCGCTCAAGAAAAGTACTTGAACC 1180

Qy  1273 AGGTTACGGGTTG-----GGGAGGGCGCACGGGTTGCTCATGGGGGTGGGGAAGT 1326
         || ||| ||                ||||||||| |||| ||||| |||| |||
Db  1181 AGATTTACGGCCAACAAACCCGGAAGGCCACACGGTTTGTTTATGGGCGGTGAGGAAAAC 1240

Qy  1327 TGCTGGCGCGCACGTGATTCAGATTCTCGTGATTGCTGGGTGGGTTAGTGCGACCATGG 1386
         | || |||| || ||||||||||| |  |||||| | ||| |||| |||||||||||||
Db  1241 TACTTGGGAGCTCAGCTGATTCAGATCATTGTGATCACGGGGTTGGGTAAGTGCGACCATGG 1300

Qy  1387 GACCCTTGTTTTGGGGGTTGAATAAACTGAAGCTGTTGAGGATTTCTTCAGAGGATGAGC 1446
         | ||| |    ||      |||   || |||||||||||| |  ||||||||||||||||
Db  1301 GGACACTTTTCTTCATCCTCAAGAAAAATGAAATTGTTGCGGATATCGTCCGGAGGATGAGA 1360

Qy  1447 TTCGCGGGATGGACATGACTCGCGCCATGGAGGCTTTGCTTATGCTTATGAGGATGATGAGA 1506
         | |||| |||||| |||||| |||| ||||| |||||||||||  ||| | | ||||| |||
Db  1361 TGGCCGGTATGGATATGACCAGGCACGGTGGTTTTGGTTTTGCTTATATGTACTTTGATGATG 1420

Qy  1507 CGCACAAGCATGGGATGGCAGTTGAGGAGGGTTGGGCCCAACGCGTCTTCCACACCCACCA 1566
         | | ||| ||| | |   |||||||||||| ||||   ||  ||  |||||| |||| |
Db  1421 AGTCTCACAAAGCCATTCAGCTTAGGAGAGTTGAGCCACGATCTCCTTCTCCTTCTTCGGTG 1480

Qy  1567 CTGAT 1571
         || ||
Db  1481 CTAAT 1485
```

RESULT 5
AAC44852
ID   AAC44852   standard; DNA; 1572 BP.
XX
AC   AAC44852;
XX
DT   18-OCT-2000 (first entry)
XX
DE   Arabidopsis thaliana DNA fragment SEQ ID NO: 44374.
XX
KW   Hybridisation assay; genetic mapping; gene expression control;
KW   protein identification; signal transduction pathway; metabolic pathway;
KW   promoter; termination sequence; ss.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-00301439.
XX
PR   25-FEB-1999; 99US-0121825P.
PR   05-MAR-1999; 99US-0123180P.
PR   09-MAR-1999; 99US-0123548P.
PR   23-MAR-1999; 99US-0125788P.
PR   25-MAR-1999; 99US-0126264P.
PR   29-MAR-1999; 99US-0126785P.
PR   01-APR-1999; 99US-0127462P.
PR   06-APR-1999; 99US-0127462P.
PR   08-APR-1999; 99US-0128234P.
PR   16-APR-1999; 99US-0128714P.
PR   19-APR-1999; 99US-0129845P.
PR   21-APR-1999; 99US-0130077P.
PR   23-APR-1999; 99US-0130449P.
PR   23-APR-1999; 99US-0130510P.
PR   28-APR-1999; 99US-0130891P.
PR   28-APR-1999; 99US-0131449P.

PR   30-APR-1999; 99US-0132048P.
PR   30-APR-1999; 99US-0132407P.
PR   04-MAY-1999; 99US-0132484P.
PR   05-MAY-1999; 99US-0132485P.
PR   06-MAY-1999; 99US-0132486P.
PR   06-MAY-1999; 99US-0132487P.
PR   07-MAY-1999; 99US-0132863P.
PR   11-MAY-1999; 99US-0134256P.
PR   14-MAY-1999; 99US-0134218P.
PR   14-MAY-1999; 99US-0134219P.
PR   14-MAY-1999; 99US-0134221P.
PR   14-MAY-1999; 99US-0134370P.
PR   18-MAY-1999; 99US-0134768P.
PR   19-MAY-1999; 99US-0134941P.
PR   20-MAY-1999; 99US-0135124P.
PR   21-MAY-1999; 99US-0135353P.
PR   24-MAY-1999; 99US-0135629P.
PR   25-MAY-1999; 99US-0136021P.
PR   27-MAY-1999; 99US-0136392P.
PR   28-MAY-1999; 99US-0136782P.
PR   01-JUN-1999; 99US-0137222P.
PR   03-JUN-1999; 99US-0137528P.
PR   04-JUN-1999; 99US-0137502P.
PR   07-JUN-1999; 99US-0137724P.
PR   08-JUN-1999; 99US-0138094P.
PR   10-JUN-1999; 99US-0138540P.
PR   10-JUN-1999; 99US-0138847P.
PR   14-JUN-1999; 99US-0139119P.
PR   16-JUN-1999; 99US-0139452P.
PR   16-JUN-1999; 99US-0139453P.
PR   17-JUN-1999; 99US-0139492P.
PR   18-JUN-1999; 99US-0139454P.
PR   18-JUN-1999; 99US-0139455P.
PR   18-JUN-1999; 99US-0139456P.
PR   18-JUN-1999; 99US-0139457P.
PR   18-JUN-1999; 99US-0139458P.
PR   18-JUN-1999; 99US-0139459P.
PR   18-JUN-1999; 99US-0139460P.
PR   18-JUN-1999; 99US-0139461P.
PR   18-JUN-1999; 99US-0139462P.
PR   18-JUN-1999; 99US-0139463P.
PR   18-JUN-1999; 99US-0139750P.
PR   18-JUN-1999; 99US-0139763P.
PR   21-JUN-1999; 99US-0139817P.
PR   22-JUN-1999; 99US-0139899P.
PR   23-JUN-1999; 99US-0140353P.
PR   23-JUN-1999; 99US-0140354P.
PR   24-JUN-1999; 99US-0140695P.
PR   28-JUN-1999; 99US-0140823P.
PR   29-JUN-1999; 99US-0140991P.
PR   30-JUN-1999; 99US-0141287P.
PR   01-JUL-1999; 99US-0141842P.
PR   01-JUL-1999; 99US-0142154P.
PR   02-JUL-1999; 99US-0142055P.
PR   06-JUL-1999; 99US-0142390P.
PR   08-JUL-1999; 99US-0142803P.
PR   09-JUL-1999; 99US-0142920P.
PR   12-JUL-1999; 99US-0142977P.
PR   13-JUL-1999; 99US-0143542P.
PR   14-JUL-1999; 99US-0143624P.
PR   15-JUL-1999; 99US-0144005P.
PR   16-JUL-1999; 99US-0144085P.
PR   16-JUL-1999; 99US-0144086P.
PR   19-JUL-1999; 99US-0144325P.
PR   19-JUL-1999; 99US-0144331P.
PR   19-JUL-1999; 99US-0144332P.
PR   19-JUL-1999; 99US-0144333P.
PR   19-JUL-1999; 99US-0144334P.
PR   19-JUL-1999; 99US-0144335P.
PR   20-JUL-1999; 99US-0144352P.
PR   20-JUL-1999; 99US-0144632P.
PR   20-JUL-1999; 99US-0144884P.
PR   21-JUL-1999; 99US-0144814P.

```
PR   21-JUL-1999;  99US-0145086P.      PR   21-OCT-1999;  99US-0160741P.
PR   21-JUL-1999;  99US-0145088P.      PR   21-OCT-1999;  99US-0160767P.
PR   22-JUL-1999;  99US-0145085P.      PR   21-OCT-1999;  99US-0160768P.
PR   22-JUL-1999;  99US-0145087P.      PR   21-OCT-1999;  99US-0160770P.
PR   22-JUL-1999;  99US-0145089P.      PR   21-OCT-1999;  99US-0160814P.
PR   22-JUL-1999;  99US-0145192P.      PR   21-OCT-1999;  99US-0160815P.
PR   23-JUL-1999;  99US-0145145P.      PR   22-OCT-1999;  99US-0160980P.
PR   23-JUL-1999;  99US-0145218P.      PR   22-OCT-1999;  99US-0160981P.
PR   23-JUL-1999;  99US-0145224P.      PR   22-OCT-1999;  99US-0160989P.
PR   26-JUL-1999;  99US-0145276P.      PR   25-OCT-1999;  99US-0161404P.
PR   27-JUL-1999;  99US-0145913P.      PR   25-OCT-1999;  99US-0161405P.
PR   27-JUL-1999;  99US-0145918P.      PR   25-OCT-1999;  99US-0161406P.
PR   27-JUL-1999;  99US-0145919P.      PR   26-OCT-1999;  99US-0161359P.
PR   28-JUL-1999;  99US-0145951P.      PR   26-OCT-1999;  99US-0161360P.
PR   02-AUG-1999;  99US-0146386P.      PR   26-OCT-1999;  99US-0161361P.
PR   02-AUG-1999;  99US-0146388P.      PR   28-OCT-1999;  99US-0161920P.
PR   02-AUG-1999;  99US-0146389P.      PR   28-OCT-1999;  99US-0161992P.
PR   03-AUG-1999;  99US-0147038P.      PR   28-OCT-1999;  99US-0161993P.
PR   04-AUG-1999;  99US-0147204P.      PR   29-OCT-1999;  99US-0162142P.
PR   04-AUG-1999;  99US-0147302P.
PR   05-AUG-1999;  99US-0147192P.
PR   05-AUG-1999;  99US-0147260P.
PR   06-AUG-1999;  99US-0147303P.
PR   06-AUG-1999;  99US-0147416P.
PR   09-AUG-1999;  99US-0147491P.
PR   09-AUG-1999;  99US-0147935P.
PR   10-AUG-1999;  99US-0148171P.
PR   11-AUG-1999;  99US-0148319P.
PR   12-AUG-1999;  99US-0148341P.
PR   13-AUG-1999;  99US-0148565P.
PR   13-AUG-1999;  99US-0148684P.
PR   16-AUG-1999;  99US-0149368P.
PR   17-AUG-1999;  99US-0149375P.
PR   18-AUG-1999;  99US-0149426P.
PR   20-AUG-1999;  99US-0149722P.
PR   20-AUG-1999;  99US-0149723P.
PR   20-AUG-1999;  99US-0149929P.
PR   23-AUG-1999;  99US-0149902P.
PR   23-AUG-1999;  99US-0149930P.
PR   25-AUG-1999;  99US-0150556P.
PR   26-AUG-1999;  99US-0150884P.
PR   27-AUG-1999;  99US-0151065P.
PR   27-AUG-1999;  99US-0151066P.
PR   30-AUG-1999;  99US-0151080P.
PR   31-AUG-1999;  99US-0151303P.
PR   01-SEP-1999;  99US-0151438P.
PR   07-SEP-1999;  99US-0151930P.
PR   07-SEP-1999;  99US-0152363P.
PR   10-SEP-1999;  99US-0153070P.
PR   13-SEP-1999;  99US-0153758P.
PR   15-SEP-1999;  99US-0154018P.
PR   16-SEP-1999;  99US-0154039P.
PR   20-SEP-1999;  99US-0154779P.
PR   22-SEP-1999;  99US-0155139P.
PR   23-SEP-1999;  99US-0155486P.
PR   24-SEP-1999;  99US-0155659P.
PR   28-SEP-1999;  99US-0156458P.
PR   29-SEP-1999;  99US-0156596P.
PR   04-OCT-1999;  99US-0157117P.
PR   05-OCT-1999;  99US-0157753P.
PR   06-OCT-1999;  99US-0157865P.
PR   07-OCT-1999;  99US-0158029P.
PR   08-OCT-1999;  99US-0158232P.
PR   12-OCT-1999;  99US-0158369P.
PR   13-OCT-1999;  99US-0159293P.
PR   13-OCT-1999;  99US-0159294P.
PR   13-OCT-1999;  99US-0159295P.
PR   14-OCT-1999;  99US-0159329P.
PR   14-OCT-1999;  99US-0159330P.
PR   14-OCT-1999;  99US-0159331P.
PR   14-OCT-1999;  99US-0159617P.
PR   14-OCT-1999;  99US-0159638P.
PR   18-OCT-1999;  99US-0159584P.
```

Query Match   44.3%;   Score 755.8; DB 3;   Length 1572;
Best Local Similarity   71.7%;   Pred. No. 6.5e-146;
Matches 1036; Conservative   0; Mismatches 397; Indels 12; Gaps   3;

```
Qy   136 ACGGCTCCGCGCCGCCTCCTTATCTGCGGCCATTCGCCGCGTGACAGCAAGTTCG 195
             |||||| |||||| |||| ||  |||| || || |||| ||| || |||
Db    83 ATGCCACGGCGGCGGCGGCCAACTACATCTGTGGCCAGTTAGGCGACGTCAACAACAAAATTTA 142

Qy   196 TCGACACGGCCTTCGGCGTCGACAACACCTACCTCTCTTTTCCGCTACCTCGTTTTTT 255
            ||||| | |||| |||  |||| ||| ||| || |||||| |||||||||||||
Db   143 TCGACACCCGCTTTCGCTATAGACAACACTTACCTTCTTCCTCGCCTTGTCTCTTCT 202

Qy   256 CTATGCAGCTCGGCTTCGGCCATGCTCTGCGCCGGTCTCGTCCGTCCGGCCAAGAACACCATGA 315
             |||||||||||||||||| ||| ||| | |||||| |||| || | || |||
Db   203 CTATGCAGCTTGGCTTCGGCTCGTGTGCCGGTTCCGTGACGGCCAAGAATACTATGA 262

Qy   316 ACATCATGCTCACCAAGTCCTGGACGACGTCCTGCGCCGGCGGCCTCTTCTACCTCTTCG 375
             |||||||| |||||||| || |||  ||||| ||||||| |||| |||||||
Db   263 ACATCATGCTTACCAAGTCCTTGACGCGTCCGCAGCCGGTCTCTTCTTCTATTATCTGTTTG 322

Qy   376 GCTTCGCCTTCGCTTCGGCTCCCCCTCCAACGGCTTCATCGGTAAACATTCTTCCGCC 435
             ||||||||||| || || ||||| || ||||||||| ||| ||||| || |||||
Db   323 GCTACGCCCTTTGCCTTTGGATCTCCGTCCAATGGTTCATCGGTAAAACTACTTTGTC 382

Qy   436 TCAAGGACATCCTTCATCTCCTACGACTACAGCTACTTCCTCTACGACGCACAGTTCGTGGCCT 495
             ||||||||||||| | |||||||| |||||| | ||| ||| ||| |||| |||||||
Db   383 TCAAAGACATCCCCACGGCCTCTGCTGACTACTCCAACTTTCTCTACCAATGGCCTTTG 442

Qy   496 CCATCGCCGCCGCGGCATCACCAGCGGAGCATCGCCGACGCACAGTTCGTGGCCT 555
             |||| |||| ||| ||||||| || ||||| ||||| ||||| ||||||||||
Db   443 CAATCGCTGCGGCTGGAATCACAAGTGGCTCGATCGCGATCGGACGACAGTTCGTGGCTT 502

Qy   556 ATCTCATCTACTCCTCCTTCCTCACCGGCTTCGTCTCATCGCTGTCTCCCACTGGTTCT 615
             ||| || | ||| ||||||||| ||| | |||||| |||||| | |||||||||||||
Db   503 ACCTAATCTATTCCTCTTTCTCTTAACCGGGTTGTTTTACCCGGTCGTCTCTCACTGGTTCT 562

Qy   616 GGTCCCCAGACGGCTGGGCCTCTGCCTT---TAAGATCACCGACCGGCTATTTTCACCG 672
             |||| || ||| | |||| | ||||| |       || |||||| || ||| |||||
Db   563 GGTCAGTTGATGGATGGGCCAGCCCGTTCCGTACCGATGGAGATTTGCTTTTCAGCACCG 622

Qy   673 GCGTAATAGACTTCGCCGTTCGGCGTAGTCCACATGGTCGGCGGAATAGCGGCCTAT 732
             |||  || |||||| ||||| ||| ||| || || || ||| |||| ||||| |||
Db   623 GAGCGATAGATTTCGCTGGTCCGGTGTGTTGTTCATATGGTCGGAGTATCGCTGGACTCT 682

Qy   733 GGGGAGCGCGTGATCGAAGGCCCAAGAAATGGGACGTTTCGATCATGCAGGACGAGCTCGTGG 792
             ||||| |||||| || ||| |  ||||| |||||| |||| || |||| |||||| ||||||
Db   683 GGGGTGCGCTCATCGAAGTCCACGACGTCCGGTTCGGCCGGTTCGATAACGGAGGCCGTGCCATCG 742

Qy   793 CCTTGCGAGGCCACAGCGCGTCCTTAGTAGTCCTGGGAACCTTCTTGCTTGTTCGGTT 852
             |||||| |||| | |||| |||| || |||| || |||| |||| ||| ||| |||||
Db   743 CTCTTCGTGGGCCACTCGGCGTCACTTGTTGTCCTTGGAACATTCCTCTCTCTGGTTTGGAT 802

Qy   853 GGTACGGGATTTAACCCCGGTTCATTTCATTTAACAAAATCCTACTTTACTTACGGTAACTCAGGAA 912
```

```
XX
PD    04-MAY-1995.
XX
PF    28-OCT-1993;        93DE-04337597.
XX
PR    28-OCT-1993;        93DE-04337597.
XX
PA    (GENB-) INST GENBIOLOGISCHE FORSCHUNG.
XX
PI    Frommer W, Ninnemann O;
XX
DR    WPI; 1995-171399/23.
DR    P-PSDB; AAR74676.
XX
PT    Plant ammonium transporter DNA in sense or anti:sense orientation -
PT    useful for transforming cells, esp. plant cells for prodn. of transgenic
PT    plants with altered nitrogen metabolism.
XX
PS    Claim 3; Page 17-20; 23pp; German.
XX
CC    The ammonium transporter (AT) gene isolated from Arabidopsis thaliana
CC    (AAQ88327) is a preferred AT gene for use in generating transgenic
CC    plants. AT gene sequences can be incorporated into expression constructs
CC    in the sense orientation (for expression of translatable mRNA and hence
CC    synthesis of AT in the transgenic plants) or in the antisense orientation
CC    (resulting in antisense transcripts which inhibit synthesis of endogenous
CC    AT in the transgenic plants). Nitrogen metabolism is altered in the
CC    resulting transgenic plants; the changes in AT activity may also result
CC    in plants which are suitable for "low input" crop management or which can
CC    be grown in acid soils. Yeast and bacteria can also be transformed with
CC    the AT gene.
XX
SQ    Sequence 1748 BP; 398 A; 427 C; 421 G; 502 T; 0 U; 0 Other;

Query Match             44.3%;  Score 755.8;  DB 2;  Length 1748;
Best Local Similarity   71.7%;  Pred. No. 6.7e-146;
Matches 1036;  Conservative    0;  Mismatches  397;  Indels  12;  Gaps   3;

Qy        136 ACGGCTTCGGCCGCCGCCTCCCTTATCTGCGGCCATTTCGCCGCGTGGACAGCAAGTTCG 195
                |  ||||||||  || || || ||||||| |||||  ||| |  ||  ||||| ||| |||
Db         64 ATGCCACGGCGGGGCGGGGCCAACTACATATGTGCCAGCTAGGCGACGTCAACAACAAATTCA 123

Qy        196 TCGAACACGGCCTTGCGCCGTGCGACAACACCTACCTGCTCTTTTCCGCTCCTACCTCGTTTTT 255
                ||||| |||||||||  || | |||| ||   |||||||||||||||||||||||||||||||
Db        124 TCGAACACGGCTTTGCGTATAGACAACACTTACCTCCTCTTCTCCGCCTACCTTGTCTCT 183

Qy        256 CTATGCAGCTCGGCTTCGCCATGCTGCGCGGCTCCGTCCGCGGCAAGAACAACCATGA 315
                ||||||||||||||||||||||||||||||| |||||| ||| ||||| || ||||||
Db        184 CTATGCAGCTTGGCTTCGCCTATGCTCGTGTGCGCCGGTTCCGTGAGAGCCAAGAATACTATGA 243

Qy        316 ACATCATGCTCACCAAGTCCTGGACGCTGCCGCCGGCGGTCGCAGCGGTGTCTCTTCTACTACCTCTTCG 375
                ||||||||||||||||||  ||||||||||||||| ||| ||| || ||| || |||  |||| ||| ||
Db        244 ACATCATGCTTACCAACGTCCTTGACGCGTGGGTTTGTTCGCTCTCTATTATCTGTTG 303

Qy        376 GCTTCGGCCTTGCTTGGCGCCCATGGAGGCTTTGCTTATGCTTATGAGGATGAGGA 435
                ||||||||||||||||||||||  || |||||||| || ||  ||| || || |||
Db        304 GCTACGGCCTTTGCTTTGGATCTCCGTCCAATGGTTTCATCGGTAAACACTACTTTGGTC 363

Qy        436 TCAAGGACATCCCTTCATCCTCCTACGACTACAGCTACTTCCTCTACCATGGGCCTTCG 495
                ||||||| || || ||  ||   |||||||||||| || ||||||  |||||||| |||
Db        364 TCAAAGACATCCCAACCCCTCTGCTGACTACTCCAACTTTCCTCTACCATGGGGCTTTG 423

Qy        496 CCATCGCGCCGCGCCGCATCACCAGCGGAAGCATCGCCGAACCGACCACACAGTTCGTTGGCCT 555
                ||||||||| |||| ||||||||||||  ||||||||||||||||||| ||||||||||||||
Db        424 CAATCGCTGCGCGCTGGAATCACAAGTGCTCAAGTGCTGCTGCAACGGACAGTTCGTGGGCTT 483

Qy        556 ATCTCATCTACTCCTCCTTCCTCACCGGTTGGTCTCCCACTGGTTCT 615
                ||| | ||  |  |||||||| ||||| |||||| ||||| |||||||
Db        484 ACCTAATCTATTCCTCTTTCTTAACCGGGTTTGTTACCCGGTCTCTCTCACTGTTCT 543

Qy        616 GGTCCCCAGAGACGGGCTGGGCCTTCGCCTT---TAAGATCACCGACCGGCTATTTTCCACCG 672
                ||||  |||||||||||||||||||||||||   ||||||||||| |||||||  || |||||
Db        544 GGTCAGTTGATGGATGGAGCCCAGCCCGTTCCGTACCGATGGAGATTTGCTTTCAGCACCG 603
```

(Second half — continuation near bottom)

```
Db        803 GGTACGGGATTTAACCCCGGTTCCTTCCTCAACAAGATCCTAGTCACGTACGACGAGA---CAGGCA 859
                ||||||||||||||||||||||||||||||||| ||||| || |  ||| ||  ||     |||||
Qy        913 ATTACTACGGTCAATGGAGCGCGGTTGGCAGAACCGCGGTCACCACTACCCTAGCGGGGT 972
                ||||  |||| || ||||||||||||| ||| |||| |||  |||||||||||| |||||
Db        860 CATACAACGGCCAGTGGACGGCGGGCGGGTCGGACGACAGCTGTCACAACGTTAGCTGGCT 919

Qy        973 CAACAGCTGCCTTGCACCAGCCAGCTATTCGGTAAACGCGGTGATATCCGGTCACTGGAACGTGA 1032
                ||||||||| ||||| ||| ||| ||||||||||   || ||  |||||||||| ||||||||||
Db        920 GCACCGCGCGCTGACAACCCTATTTGGAAACGTCACTCTCGGGACATTGGAACGTCA 979

Qy       1033 CCGATGTCTGCAACGGGCTGTTAGGCGGTTTCGCGCCGATAACAGCCGGTTGCTCCGTGG 1092
                |||||| ||||||||||||||| ||||| | ||||||||| |  |||| ||||| |||||
Db        980 CTGATGTATGCAACGACGCCCTCCCTCGGAGGGTTTGCAGCAGTCGCCATAACTGCTCCTCTCGTCG 1039

Qy       1093 TTGAGCGCCATGGGCAGCCATCGTATGCGGTTTTGTTGCTTCTATAGTATTAATAGCTTGCA 1152
                ||||||||||||| ||| || ||| ||| |||| ||| || |||  ||  |  || |||||
Db       1040 TTCAGCCATGGCCTGCGATCATCTGCGGGTTCGTGCGCCCTAGTCCTCCTCGGATGCA 1099

Qy       1153 ACAAATTAGCAGAGAAGGTTAAGTTCGACGATCCTCTGGAGGCGCGCGCGAGTTGCACGGTG 1212
                ||||| || ||||||| || ||||||||||||||||||| |||||| || ||||| | ||||
Db       1100 ACAAGCTCGCTGAGAAGTCAAATACGACGACCCCTTGGAGGCAGCACAACTACACGGTG 1159

Qy       1213 GGTGTGGCACGTGGGGGGTGATATTCACCGCGTTGTTCGCAAAAAGGAGTATGTGAAGG 1272
                |||||||||| || |||  ||| | |||| | ||| |||| ||||| |||  || |||||
Db       1160 GTTGCGGTGCCTGGGGACTAATATTCACCGGCTCTCGTTCCGCTCAAGAAAAGTACTTGAACC 1219

Qy       1273 AGGTTTACGGGTG------GGGAGGGGCCACGGGTTGCTCATGGGGGTGGTGGGAAGT 1326
                ||| |||||  ||      |||| ||||| | ||||  | |||  | || |||||||
Db       1220 AGATTTACGGCACAACAACCCGGAAGGCCACCGGGTTTGTTTATGGCGGTGGAGGAAAAC 1279

Qy       1327 TGCTGGGGCGGCCACGTGATTCGTGGTGATTGCTGGGTGGTTAGTGCGACCATGG 1386
                ||| | ||| ||| ||| || || || ||||| |  |||| |||||||| ||||||
Db       1280 TACTTGGAGCTCAGCTGATTCAGATCATTGTGATCACGGGTTGGGTAAGTGCGACCATGG 1339

Qy       1387 GACCCTTGTTTGGGGGGTTGAATAAACTGAAGCTGTTGAGGATTTCTTCAGAGGATGAGC 1446
                ||  ||| |||| |||||||||| || ||||| ||| || || || ||| ||||||||||
Db       1340 GGACACTTTTCTTCATCCTCAAGAAAATGAAATTGTTGCGGATATCGTCCGAGGATGAGA 1399

Qy       1447 TTGCGGGGATGGACAATGACTCGCCATGGAGGCGTTTGCTTATGCTTATGAGGATGATGAGA 1506
                ||| |||||| || |||||||||| |||| ||| |||||| ||  |||||| || |||||||
Db       1400 TGGCCGTATGATGTAGACCAGCGACGGTGGTTTTTGCTTATATATGTTATTTGATGATGATG 1459

Qy       1507 CGCACAAGCATGGGATGCAGTGAGGAGGGTTGGGCCCAACGCGTCTTCCACACCCACCA 1566
                |||||||||| || |  ||| |||||| ||||||||||| || ||||||||  ||||||
Db       1460 AGTCTCACAAAGCCATTCAGCTTAGGAGAGTTGAGCCATCTTCTTCCGACTACTTCCAACTTTCCTCTG 1519

Qy       1567 CTGAT 1571
                ||||
Db       1520 CTAAT 1524
```

```
RESULT 6
AAQ88327
ID   AAQ88327 standard; cDNA; 1748 BP.
XX
AC   AAQ88327;
XX
DT   05-JAN-1996 (first entry)
XX
DE   Arabidopsis thaliana ammonium transporter gene.
XX
KW   Nitrogen metabolism; transgenic plant; ammonium transporter gene;
KW   antisense; inhibition; low input crop management; ds.
XX
OS   Arabidopsis thaliana.
XX
FH   Key            Location/Qualifiers
FT   CDS            21..1526
FT                  /*tag= a
FT                  /product= "Ammonium_transporter"
XX
DR   DE433597-A1.
```

```
XX
DE   Arabidopsis thaliana stress regulated gene SEQ ID NO 1490.
XX
KW   Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
XX
OS   Arabidopsis thaliana.
XX
PN   WO200216655-A2.
XX
PD   28-FEB-2002.
XX
PF   24-AUG-2001; 2001WO-US026685.
XX
PR   24-AUG-2000; 2000US-0227866P.
PR   26-JAN-2001; 2001US-0264647P.
PR   22-JUN-2001; 2001US-0301111P.
XX
PA   (SCRI ) SCRIPPS RES INST.
PA   (SYGN ) SYNGNTA PARTICIPATIONS AG.
XX
PI   Harper JF, Kreps J, Wang X, Zhu T;
XX
DR   WPI; 2002-304127/34.
XX
PT   Identifying a stress condition to which a plant cell has been exposed and
PT   producing plants with increased tolerance to these abiotic stresses.
XX
PS   Claim 144; SEQ ID NO 1490; 577pp + Sequence Listing; English.
XX
CC   The invention relates to identifying a stress condition to which a plant
CC   cell has been exposed, comprising: (a) contacting nucleic acid
CC   representative of expressed polynucleotides in the plant cell with an
CC   array or probes representative of the plant cell genome; and (b)
CC   detecting a profile of expressed polynucleotides in the plant cell
CC   characteristic of a stress response. The method is useful in the
CC   production of transgenic plants, cells and seeds and in producing plants
CC   with increased tolerance to abiotic stress. The present sequence is that
CC   of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used
CC   in methods of the invention. Note: The sequence data for this patent is
CC   not represented in the printed specification but is based on sequence
CC   information supplied to Derwent by the European Patent Office
XX
SQ   Sequence 1545 BP; 326 A; 392 C; 382 G; 445 T; 0 U; 0 Other;
```

```
Query Match         41.4%;   Score 705.8;   DB 6;   Length 1545;
Best Local Similarity  70.3%;     Pred. No. 1.3e-135;
Matches 979;    Conservative   0;   Mismatches 402;   Indels 12;   Gaps   2;
```

```
Qy   673  GGGTAATAGACTTCGCCGGTTCCGGCGTAGTCCACATGGTCGGCGGAATAGCCGGCCTAT  732
               | |||||| ||||||  ||||  |  || |||||||||||||||
Db   604  GAGCGATAGATTTCGCTGGGTCCGGTGTTGTTCATATGGTCGGAGTATCCGTGGACTCT   663

Qy   733  GGGGAGAGCGCTGATCGAAGAATGGACGCCCAAGAATGGACGTTTCGATCATCGGAGCGAGCTGTGG  792
               |||| |||  ||||||||||  |||  ||  ||||||||||   |||
Db   664  GGGGTGCGCTCATCGAAGGTCCACGACTTGGCCGGTTCGATAAACGGAGGGCCGTGCCATCG   723

Qy   793  CCTTGCGAGGCCACAGCGCGTCCTTAGTAGTCCTGGGAACCTTCTTGCTTGGTTCGGTT   852
               |||| ||  |||||| |||||  ||||  ||||||||||||||||||
Db   724  CTCTTCGTGCGGCCACTCGTTGTTCCTTGGAACATTCCTCCTCTTGGTTTGGAT   783

Qy   853  GGTACGGATTTAACCCCGGTTCATTTAACAAAATCCTACTTACTTACGGTAACTCAGGAA  912
          ||||||||||||||||||||||  ||||||| |||||| ||  |||||  ||| ||||||
Db   784  GGTACGGATTTAACCCCGGTTCCTTCAACAAGATCCTAGTCACGTACGAGA--CAGGCA   840

Qy   913  ATTACTACGGTCAATGGAGCGCGTTGGCAGAACCGCGGTCACCACTACCCTAGGGGGGT  972
               |||| |||| |||||||||||  || ||| |||| ||||||||||
Db   841  CATCACAACGGCCAGTGAGCGCGTCGGACGGACGAGCGTGTCACAACAACGTTAGCTGGCT   900

Qy   973  CAACAGCTGCCTTGACCACGCTATTCGGTAAACGGGTGATATCCGTCACTGGAACGTGA  1032
               | |||| ||||| ||  |||||||||  |||||  || ||||||||||||
Db   901  GCACCGCGCGCTGACAACCCTATTGGGAAACGTCTACTCCTCGGGACATTGGAACGTCA   960

Qy  1033  CCGATGTCTGCAACGGCTGTTAGGCGGTTTCGCGCGGATAACAGCCGGTTGCTCCGTGG  1092
               |||||| ||||||||  ||| ||| |||  || ||||||| |||||||||| |
Db   961  CTGATGTATGCAACGGCCCTCCTCGGAGCGGTTTGCAGCCATAACTGTGGCTGCTCTGTCG  1020

Qy  1093  TTGAGCCATGGGCAGCCATCGTATGCGGTTTTGTTGCTTCTATAGTATTAATAGCTTGCA  1152
          |||||||| |||| ||| ||||  ||||||| ||||| |||||
Db  1021  TTGAGCCATGGCCAGATCATCGCGGTTCGTGCGCGCCCAGTCCTTCCTCGGATGCA  1080

Qy  1153  ACAAATTAGCAGAGAAGGTTAAGTTCGACGATCCTCTGGAGGCGCGCAGTTGCACGGTG  1212
               ||  ||||||||||| ||||| |||||||||| ||||||  ||||||||||
Db  1081  ACAAGCTCGCTGAGAAGTCAAATACGACGACCCTCTTGAGGCAGCAGCACACACGGTG  1140

Qy  1213  GGTGTGCACGTGGGGGTGATATTCACGGCGTTGTTCGCAAAAAAGAGTATGTGAAGG  1272
               ||||| ||||||  ||| || ||||  | |||||   ||||||||||
Db  1141  GTTGCGGTGCGTGGGGACTAATATTCACGGCTCTTCTTGCTCAAGAAAAGTACTTGAACC  1200

Qy  1273  AGGTTTACGGGTTG------GGGAGGGCGCACAGGTTGCTCATGGGGTGGTGGGAAGT  1326
          || |||||||  ||       ||||| ||| ||||||  ||  |||| ||||| ||||
Db  1201  AGATTTACGGCAACAAACCCGGAAGGCCACACGGTTTGTTTTATGGGCGGTGAGGAAAAC  1260

Qy  1327  TGCTGGGGGCGCACGTGATTCAGATTCTGGTGATTGCTGGGTGGGTTAGTGCGACCATGG  1386
               |||| |||||||| |||||||||  |||| || || ||||||| ||||||||||||
Db  1261  TACTTGGAGCTTCAGCTTGATTCAGATCATTGTGATCACGGGTTGGGTAAGTGCGACCATGG  1320

Qy  1387  GACCCTTGTTTTGGGGGTTGAATAAACTGAAGCTGTTGAGGATTTCTTCAGAGGATGAGC  1446
               || |||| ||||||||| |||| ||||| |||||||| ||||| |||||
Db  1321  GGACACTTTTCTTCATCCTCAAGAAAAATGAAATTGTTGCGGATATCGTCCAGGAGTGAGA  1380

Qy  1447  TTGCGGGGATGGACATGACTCGCCATGGAGGGCTTTGCTTATGGTTATGAGGATGGATGAGA  1506
          | || |||| | |||||||  || ||  ||| ||||||||||  ||| ||||| |||||
Db  1381  TGGCCGGTATGATATGACCAGGCACGGTGGTTTGCTTATATGTACTTTGATGATGATG  1440

Qy  1507  CGCACAAGCATGGGATGGATGCAGTTGAGGAGGGGTTGGGCCCAAGCGCGTCTTCCACACCCACCA  1566
               ||||||| | || |||  |  |||||||  ||||| || |  || ||||
Db  1441  AGTCTCACAAAGCCATTCAGCTTAGGGAGAGTTGAGCCACGAGTCCAGCACGATCTCCTTCTCCTTCTCGGTG  1500

Qy  1567  CTGAT  1571
          || ||
Db  1501  CTAAT  1505
```

```
Qy   124  CAAACACCACAGACGCGCTCCCGCCCGCCTCCTTATCTGCGGCCATTCGCCGCCGCTGG  183
               | ||| ||| || |||  ||||||| ||||| |||||||||
Db    62  CTAACTCAACATCTTCCCTCGCCGCGGCAACCTTTTTATGTTCCCAAATTTCAAACATCT  121

Qy   184  ACAGCAAGTCGTCGACACGGCCTTCGCCGTGACAACAACTACCTCCTCTTTTCCGCCT  243
               | ||||| | | |||||| || ||| ||| |||||||||||||||||  |||||
Db   122  CCAACAAACTCTCCGACACCAACTTATGCCGTCGACAACACGTATCCTCCTCTCTCGCCT  181

Qy   244  ACCTCGTTTTTCTTATGCAGCTCGGCTTCGCCATGCTTCGCGCCGCCATGCTTCGTCCGGTCCGTCCGGCGCCA  303
          ||||  |  ||| | ||||| |||| |||| |||||||  | ||  || |||||  ||||  |||  || ||||
Db   182  ACCTTGTCTTTGCCATGCAGTCGGTTTCGTCTATGCTTTGTCTGATCAGTCCGAGCCA  241

Qy   304  AGAACACCACATGAACAATCATGCTCCACCACGTCCTGGACGGTCGCCGCGGGCCTCTTCT  363
          ||||||| | |||||| ||||||| | || |||||  || || || |  | | || ||||
Db   242  AGAACACTATGAACATCATGCTTACCAATGTCCTTGATGCTGCCGGACCATCTCTT  301

Qy   364  ACTACCTCTTCGGCTTCGCCTTCGCCTTTCGGCTCCCCCTCCAACGGCTTCATCGGTAAAC  423
          |||||||||||| |||||| ||||||||| || || ||||| ||||| |||||||| | |
Db   302  ACTACCTCTTCGGATTCGCATTCGCCTTTGGTACACCTTCCAACGGATTCATCGGTCGCC  361

Qy   424  ATTTCTTCGGGCCTCAAGGACATCCCTT-------CATCCTCCTACGGACTACAGCTACT  474
               || || ||||  | ||  |||            ||  | |||  ||||  |||  |
Db   362  ACCATAGCTTCTTCGCTTTAAGCTTCTTACCCTGAACGCCCCCGGCTTCCGACTTCAGTTTT  421

Qy   475  TCCTCTACCACCAATGGGCCTTCGCCATGCATCGCCGGCCGCGGCGCCGGCATCACCAGGCGGAAGCATCGCCG  534
```

```
RESULT 7
AB213685

ID   AB213685   standard; DNA; 1545 BP.
XX
AC   AB213685;
XX
DT   21-JAN-2003   (first entry)
```

```
ABZ14082
ID   ABZ14082 standard; DNA; 1497 BP.
XX
AC   ABZ14082;
XX
DT   21-JAN-2003  (first entry)
XX
DE   Arabidopsis thaliana stress regulated gene SEQ ID NO 1887.
XX
KW   Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
XX
OS   Arabidopsis thaliana.
XX
PN   WO200216655-A2.
XX
PD   28-FEB-2002.
XX
PF   24-AUG-2001; 2001WO-US026685.
XX
PR   24-AUG-2000; 2000US-0227866P.
PR   26-JAN-2001; 2001US-0264647P.
PR   22-JUN-2001; 2001US-0300111P.
XX
PA   (SCRI ) SCRIPPS RES INST.
PA   (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX
PI   Harper JF, Kreps J, Wang X, Zhu T;
XX
DR   WPI; 2002-304127/34.
XX
PT   Identifying a stress condition to which a plant cell has been exposed and
PT   producing plants with increased tolerance to these abiotic stresses.
XX
PS   Claim 144; SEQ ID NO 1887; 577pp + Sequence Listing; English.
XX
CC   The invention relates to identifying a stress condition to which a plant
CC   cell has been exposed, comprising: (a) contacting nucleic acid
CC   representative of expressed polynucleotides in the plant cell with an
CC   array or probes representative of the plant cell genome; and (b)
CC   detecting a profile of expressed polynucleotides in the plant cell
CC   characteristic of a stress response. The method is useful in the
CC   production of transgenic plants, cells and seeds and in producing plants
CC   with increased tolerance to abiotic stress. The present sequence is that
CC   of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used
CC   in methods of the invention. Note: The sequence data for this patent is
CC   not represented in the printed specification but is based on sequence
CC   information supplied to Derwent by the European Patent Office
XX
SQ   Sequence 1497 BP; 309 A; 386 C; 385 G; 417 T; 0 U; 0 Other;

Query Match             41.1%;  Score 700.4;  DB 6;  Length 1497;
Best Local Similarity   69.7%;  Pred. No. 1.6e-134;
Matches  995;  Conservative   0;  Mismatches  421;  Indels  12;  Gaps   3;
```

```
Qy        107 GGCCAACTTCTCGGCCCAAACACCAGACGCCTCCGCGCCGCCTCCCTTATCTGCGG 166
                ||||| ||||| ||| |||||||||||||| || || ||||  ||||||||||||
Db         27 GGCCGATCTCGGCCACCCTACTTGGCCCCAACGCCACGGCGCCGGACTACATTGCGG 86

Qy        167 CCATTTCGCCGCCGGTGGACAGCAGAAGTTCGTCGACAACGGCCTTGCCGTCGACACCACCTA 226
                |||| ||||| ||||||| |||| || || || || || ||| ||| ||  |||||||||
Db         87 CCAATTAGGCACCGTTAACAACAAGTTCACCGATGCAGCTTCCCCATAGACAACACCTA 146

Qy        227 CCTCCCTCTTTCCGCCTCACCTCGTTTTTCTATGCAGCTCGGCTTCGGCCATGCTCGGCGCC 286
                |||||| |||| ||  ||||| |||||||| |||||| |||| ||| ||||||||| |||||
Db        147 CCTCCCTCTTCTCTGCCTCACCTTGTCTTCGCCATGCAGCTCGGCTCGTTGTGC 206

Qy        287 CGGCTCCGTCCCGGCCCAAGAACACCATGAACATCATGCTCACCAAGCTCCTGGACGCTGC 346
                |||||| |||| ||||||||||||||||| |||| ||||| ||||||||||||||||||
Db        207 TGGTTCTGTTAGAGCCAAGAAATACGATGAACATCATGCTTACCAATGTCCTTGACGCTGC 266

Qy        347 CGCGGCGGCCTTCTTCTACTACCTCTTCGGCCTTTCGGCTCCCCTCCAA 406
                ||||| || ||| || |||| ||||||||||||||| |||| ||||||
Db        267 AGCCGGAGGACTCTTCTACTATCTTTTTGGTTACGCCCTTTGCCTTTGGAGGATCCTCCGA 326
```

```
Db         ||||||| ||  |||||| ||||  ||||              |||||
          422 TCCTTACCAATGGGCTTTTGCCATAGCCGCGGCCGGAAGTCACTACGCGGTTCCATCGCCG 481

Qy        535 AACGGCACACAGTTCGTGGCCTATCCATCACTCCTCCTCCGGTCGTCGTCTCTATC 594
                ||||| ||    |||| |||||||||||  ||| |||||| |   ||| |
Db        482 AGCCAACGCAATTCGTTGCTTACCTTATCTACTTTCTTGACCGGTTTTGTTTACC 541

Qy        595 CGGTGGTCTCCCACTGGTTCTGGTTCCCCAGACGGCTGGGC---CTCTGCCTTTAAGATCA 651
                |||||||   || ||||| ||| |||| |||| |||||||     |  |||| |||| ||
Db        542 CGACGTCCGCACTGGTTCTCGGTCAAGTGATGATGGCTAGCCGCGTCCCGGTCTGACA 601

Qy        652 CCGACCGGCTATTTCCACCGGCGTAATAGACTTCGCCGGTTCCGGCTAGTCCACATGG 711
                ||| ||| |||| ||| ||||| | || |||| ||| ||| |||| |||  ||||||
Db        602 ACAATCTCTTGTTTGGCTCAGGTGCTATTCATTTCGCAGGTTCAGCAGGTGTTCACATGG 661

Qy        712 TCGGCGGAATAGGCCGGCCCATAGGGAGCGCGTGATCGAAGGCCCAAGAAATGGGACGTTTCG 771
                ||| |||   || |||| |||| |||| ||   |||||||  |||||||| |||| |||||
Db        662 TAGGTGGAATTGCCCGGTTATGTGGAGCGGTAGTTGAAGGACCCAAGAATAGGTAGATTTG 721

Qy        772 ATCATGCAGGACGAGCTGTGGCCTTGCGAGGCCACAGCGCGTCCTTAGTAGTCCTGGGAA 831
                ||||||| || || |||| ||||| ||||| ||||||| ||| |||| |||||||| |||
Db        722 ACCGGTCAGGCCGGTCCGTTACGTGGTCACAGTGCATCCCCTTGTCGTGCTTGGTA 781

Qy        832 CCTTCTTGCTTTGGTTCGGTTGGGTACGGATTTAACCCCGGTTCATTTAACAAAATCCTAC 891
                |||  ||||| ||| ||| |||||| || |||||  |||||||||| |||||||||||
Db        782 CCTTCTTGTGTGGTTTGGATGGGTATGGGTTTAACCCTGGTTCCTTTTTAACCATTCTTA 841

Qy        892 TTACTTACGGTAACTCAGGAAATTACTACGGTCATGGAGCGCGGTTGGCAGAACCGCGG 951
                || || ||||||| | || ||||||| |||  ||||||||  |||| |||||||| |||
Db        842 AAGGCTACGACAAGTCTCGCCATATTAATGGTCAATGGAGCGCTGTAGGTCGCACCGCGG 901

Qy        952 TCACCACTACCCTAGCGGGGTCAACAGCTGCCTTGACCACGCTATTCGGTAAACGGGTGA 1011
                |||||| |||||||| ||  || |||||||||||| |||| | ||||||||||||||
Db        902 TCACCACACGCTTTCTGCGTGCACCGCTGCGTTGACTACTACTCTATTCAGTAAACGGCTTT 961

Qy       1012 TATCCGGTCACTGGAACGTGACCGATGTCTGCACAACGGGCGTGTTAGGCGGTTCGGCGGCGA 1071
                ||||| || ||||| ||||| || || |||||||| |||| |||||| |||| ||| ||
Db        962 TAGCAGGTCATTGGAACGTTATTGACGTATGCAACGGACTTCTAGGCGCTTGCAGCTA 1021

Qy       1072 TAACAGCCCGTTGCTCCGTGGTTGAGCCATGGCATGGCCATCGTATGCGGTTTTGTTGCTT 1131
                ||| | |  |||| ||| |||||| |||| |||||  | |||| ||||| |||||| |||
Db       1022 TAACCTCCGGATGTGCCGTGGTGGAGCCGGTGGCTTGGTCATAGTAGTATGTGGCTTTGTTGTGGCAT 1081

Qy       1132 CTATAGTATTAATAGCTTGCAACAAATTAGCAGAGAAGGTTAAGTTCGACGATCCTCTGG 1191
                | || || |||| |  ||| || |||||||| |||| ||||||| |||  ||| |||
Db       1082 CATGGGTTTTAATCGGATTTAACTTGCTTGCCAAGAAACTTAAAATGATGACCCCACTCG 1141

Qy       1192 AGGCGGCGCAGTTGCACGGTGGGTGTGGCACGTGGGGGGTGATATTCACCGGCGTGTTCG 1251
                |||| ||| |  |||| |||||||||| ||  || ||| |  |||||||||| |||||||
Db       1142 AGGCTGCTCAGCTCCACGGTGGATGTGGAGCATGGGAGGTGATTAATCTTTACCGGCGTTCG 1201

Qy       1252 CAAAAAGGAGTATGTGAAGGAGGTTACGGTTGGGGAGGCGGCCACGGGGTTGCTCATGG 1311
                ||||  |||||||  ||||||| ||  || || |||| |||| |  ||||||||||||
Db       1202 CAAGGAAAGAATACGCTTAACGAGATTTACTCCGGTGATAGGCCTTACGGACTGTTCATGG 1261

Qy       1312 GGGGTGGTGGGGAAGTTGCTGGCGCCGGCACGTGATTCAGATTCTCGTGATTGCTGGGTGGG 1371
                ||| ||||| | |||| |||| ||  |||||| |||| ||||||||| |||||||||||
Db       1262 GCGGGGGAGCAGAAAACTGCTCGCCGGCAGATCGTTCAGATTATTGTGATCGTTGGGTGGG 1321

Qy       1372 TTAGTGCGACCATGGGACCCTTGTTTGTGGGGTTGAATAAACTGAAGCTGTTGAGGATTT 1431
                |||| ||| ||||||  || ||| | ||| ||||||| || |||| | ||  |||| ||
Db       1322 TGACGGTAACTATGGGACCGTGTGTTTTATGGGTTACATAAGATGAATCTTTTGAGGATAT 1381

Qy       1432 CTTCAGAGGATGAGCTTGCCGGCGATGGACAATGACCTGCCCATGGAGGCTTTGCTTATGCTT 1491
                ||  ||| |||||| ||| |||| |||| | | |||| ||||  ||||||| |||| |||
Db       1382 CAGCAGAAGATGAATGGCAGCAGGAATGGCAATGGACACGTCATGGGAGAGGATTTGCTTACGCAT 1441

Qy       1492 ATGAGGATGATGA 1504
                || |  ||||| |
Db       1442 ACAATGACGAAGA 1454
```

ESULT 8

Qy   1478 CTTTGCTTATGCTTATGAGGATGATGAGAGACGCACAAGCATGGGATGCA 1525
                  ||||||||||||||||| ||||||| |||| |||||| ||||||
Db   1404 CTTTGCTTATATCTACCATGATAATGATGAGTGTCTCATAGAGTGGA 1451

RESULT 9
AAL53984
ID   AAL53984 standard; cDNA; 1037 BP.
XX
AC   AAL53984;
XX
DT   18-FEB-2003 (first entry)
XX
DE   cDNA encoding a corn high affinity ammonium transporter protein.
XX
KW   Herbicide; ammonium transporter protein; herbicide; transgenic plant;
KW   corn; gene; ss.
XX
OS   Zea mays.
XX
FH   Key   Location/Qualifiers
FT   CDS   3..554
FT   /tag= a
FT   /product= "Corn high affinity ammonium transporter
FT   protein"
XX
PN   US2002142390-A1.
XX
PD   03-OCT-2002.
XX
PF   28-DEC-2001; 2001US-00033109.
XX
PR   98US-0098248P;
PR   27-AUG-1999;   99US-00384625.
XX
PA   (ALLE/) ALLEN S M.
PA   (RAFA/) RAFALSKI J A.
XX
PI   Allen SM, Rafalski JA;
XX
DR   WPI; 2003-102520/09.
DR   P-PSDB; AAO26532.
XX
PT   Novel ammonium transporter polypeptide useful for identifying enzymatic
PT   inhibitors, which is homologous to corn, soybean, wheat or rice ammonium
PT   transporter polypeptides.
XX
PS   Claim 3; Page 12-13; 27pp; English.
XX
CC   The invention relates to a novel ammonium transporter protein comprising
CC   90% homology based on the clustal method compared to: a corn ammonium
CC   transporter polypeptide of 183 or 63 amino acids; a soybean ammonium
CC   transporter polypeptide of 500 or 486 amino acids; a wheat ammonium
CC   transporter polypeptide of 494 or 470 amino acids; and a rice ammonium
CC   transporter 497 amino acids fully defined in the specification. The
CC   isolated polynucleotide is useful for selecting an isolated
CC   polynucleotide that affects the level of expression of the ammonium
CC   transporter polypeptide in a plant cell. The ammonium transporter protein
CC   is useful for preparing antibodies which are useful for detecting the
CC   transporter protein in situ in cells or in vitro in cell extracts, and as
CC   targets to facilitate design and/or identify inhibitors of the enzymes
CC   that are useful as herbicides. The isolated polynucleotide is also useful
CC   for creating transgenic plants in which the polynucleotide is present at
CC   higher or lower levels than normal, and for designing and producing
CC   primer pairs which are useful in amplification or primer extension
CC   reactions. This polynucleotide sequence represents the cDNA encoding a
CC   corn high affinity ammonium transporter protein of the invention
XX
SQ   Sequence 1037 BP; 206 A; 250 C; 350 G; 231 T; 0 U; 0 Other;

Query Match   13.0%; Score 221; DB 7; Length 1037;
Best Local Similarity 68.5%; Pred. No.7.4e-36;
Matches 305; Conservative 0; Mismatches 140; Indels 0; Gaps 0;

Qy   407 CGGCTTCATCGGTAAACATTTCTTCGGCCTCAAGGACATCCTTCATCCTCCTACGACTA 466
                ||||||||||| | |||||||||| || | ||||||||
Db   327 AGGGTTCATTGGAAGACACAACTTGCTCTTAGAGACTTCCGACTCCACAGCTGATTA 386

Qy   467 CAGCTACTTCCTTCCTACCAAATGGGCCTTCGCCATCGCCGGCATCACCAGCGGGAAG 526
                | | ||||||||| |||| ||||||||| |||||| | ||||
Db   387 CTCTTTCTTCCTCTACCAAATGGGCGTTCGCAATCGCGGCCGCGTCGAATCACAAGTGGTTC 446

Qy   527 CATCGCGAACGCACACAGTTCGTGGCCTATCTCCATCTACTCCTCCTCACCGGCCTT 586
                ||||||||| |||||| | ||| |||||| | |||||||||||||
Db   447 GATCCGCAGAGAGGACTCAGTTCGTGGCTTACTTGATATACTCTTTCTTTCTTAACCGGATT 506

Qy   587 CGTCTATCCGGTGGTCTCCCACTGGTTCGTCCCCAGACGGCTGGGCCTCTGCCTTTAA 646
                || |||||||||||||| |||||| || | |||| || |||||
Db   507 TGTTTACCCGGTTGTCTCTCACTGGTTTGTCCCCGATGAGGGCCAGTCCCTTTCG 566

Qy   647 GATCAC---CGACCGGCTATTTTCCACGGCGTAATAGACTTGCCGGTTCCGCGTAGT 703
                |||||| || || |||| || ||| | |||||| ||
Db   567 TTCAGCGGATGATCGTTTGTTTTAGCACCGGAGCCATTGACTTTGCTCGGCTCCGGTTGT 626

Qy   704 CCACATGGTCGGCGGAATAGCCGCCTATGGGAGCGGCTGACGGGCCCAGAAGATGG 763
                |||||||| || |||| ||| | | | |||| | ||| ||||
Db   627 TCACATGGTTGGTGGCATAGCAGGTTTATGGGGTGCTCTTATTGAAGGTCCTCGTCGTGG 686

Qy   764 ACGTTTCGATCATGCAGGACGAGCTGTGCCTTGCGAGGCCACAGCGCGTCCTTAGTAGT 823
                ||||||||| | || ||||||| ||||||| || ||| ||||||||||||
Db   687 TCGGTTCGAGAAAGGTGGTCGCGCTATTGCTCGCGGGGGCCACTCTGCTCGCTAGTAGT 746

Qy   824 CCTGGGAACCTTCTTGCTTTTGGTTCGGTTGGTACGGATTTAACCCCGGTTCATTTAACAA 883
                | || |||||||| ||||||||||| | ||||||||
Db   747 CTTAGGAACCTTCCTCCTCCTATGGTTTGATGGGTATGGTTTCAACCCCGGTTCCTTCACTAA 806

Qy   884 AATCCTACTTACTTACGGTAACTCAGGAAAATTACTACGGTAAGAGGCGGGCGGTTGGCAG 943
                | | ||| | |||||||||||||||||||| || ||||
Db   807 GATACTCGTTCCGTA---TAATTCTGTTCCAACTACAGGCCAATGGAGCGGAATCGGCCG 863

Qy   944 AACCCGCGGTCACCACTACCCTAGCGGGGTCAACAGCTGCCTTGACCACGCTATTCGGTAA 1003
                ||||||| |||| |||| || || || ||||| | | |||||||||| ||||
Db   864 TACAGCGGTTAACACCACACTCCAGGATGCACAGCAGTCTAACCACTCTTTGGTAA 923

Qy   1004 ACGGGGTGATATCCGGTCACTCGAACGTGACCGATGTCTGCAACGGGCGTGTTAGGCGGTTT 1063
                | | || ||||||||||||| ||||| ||||| | ||| ||||| |||
Db   924 ACGTCTCCTATCAGGCCACTGGAACGTAACGACGTTTGCAACGGGTTACTCGTGGGTT 983

Qy   1064 CGCGGCGGATAACAGCGCGGTTGTCTCCGGTTGCCGGTTGAGCCATGGGCAGCCATCGTATGCGTTTT 1123
                |||| ||| | | | |||||||||||| | | ||||||||||||| | ||||
Db   984 TGCGGCCATAACCGCAGGTTGCTCCTCCGTAGAGCCCATGGGCAGCCATGGTGCGGCTT 1043

Qy   1124 TGTTGCTTCTATAGTATTAATAGCTTGCAACAAATTAGCAGAGAAGGTTAAGTTCGACGA 1183
                |||| || ||||| | |||| ||||||||||||||| |||||||| |||||| |||
Db   1044 CATGCGTTCTGTCCTTATCCGTATTCGGAATGCAACAAGCTCGCGGAGCTTGTACAATATGATGA 1103

Qy   1184 TCCTCTGGAGGCGGCGGCAGTTGCACGGTGGCGTGCGACGTGGGGTGATATTCACGGC 1243
                || ||| || || | | ||| |||| || |||||||| || ||||||||
Db   1104 TCCACTCGAGCGCCAACTACATGGGGTGGCGTGGGCGTGGGTTGATATTCGTAGG 1163

Qy   1244 GTTGTTCGCAAAAAAGGAGTATGTGAAGGAGGTTACGG------GTTGGGGAGGGCGCA 1297
                ||||||||||| | | | |||| | | |||| || || |||| ||
Db   1164 ATTGTTTGCCAAAGAAGAAGTATCTAAACGAGGTTTATGGCCGCCACCCCGGGAAGGCCATA 1223

Qy   1298 CGGGTTGCTCATGGGGGGTGTGGGAAGTTGCTGGCCGCCGCACGTGATTCAGATTCTGT 1357
                ||||||||| |||| | ||||| ||| | |||| | | | ||| | || ||
Db   1224 TGGACTATTTATGGGCGGAGAGGGAAGCTGTTGGGAGCACAATTGGTTCAAATACTTGT 1283

Qy   1358 GATTGCTGGGTGGGTTAGTGCGGACCATGGGACCCTTGTTTTGGGGGTTGAATAAACTGAA 1417
                || |||||| | || | || |||| || |||| || | ||||||
Db   1284 GATTGTAGGGTTAGTGCCACAATGGGAACACTCTTCTTCATCCTCAAAAGGCTCAA 1343

Qy   1418 GCTGTTGAGGATTTCTTCAGAGGATGAGCTTGCGGGATGACATGGACCATGGAGG 1477
                || |||||| || || | ||||||| | ||| | || ||| ||
Db   1344 TCTGCTTAGGATCTCGGAGCAGCAGCATGAAATGCAAGGGATGATATGACACGTCACGGTGG 1403

```
Qy   1060 GTTTCGCGGCGATAACAGCCGGTTGCTCCGTGGTTGAGCCATGGGCAGCCATCGTATGCG 1119
          |||||||||||| |||| |||||||||||||||||||||||||||||||| ||||||||
Db      7 GGTTCGCGGGCCATCACGCGCCGGGTGCAGCGTGGTGGAGCCGGTGGCGGCCGTCATCTGCG 66

Qy   1120 GTTTTGTTGCTTCTATAGTATTAATAGCTTGCAACAAATTAGCAGAGAAGGTTAAGTTCG 1179
          ||| |||||||                  |||||  |||| |||    |||   ||||||
Db     67 GGTTCGTGTCCGCGTGGGTGCTCATCGGCGCCAACGCCCTCGCGGCCCTTCAGGTTCG 126

Qy   1180 ACGATCCTCTGGAGGCCGCAGTTGCACGGTGGTGTGGCACGTGGGGGTGATATTCA 1239
          |||| |||||||||||||||||||||||||||||||||||| |||  |||||||||
Db    127 ACGACCCGCTGGAGGCCGCAGCTGCACAGCGGCGGTGTGGCGCCTGGGGCGTCCTCTTCA 186

Qy   1240 CGGCGGTTGTTGCGCAAAAAAGGAGTATGTGAAGGAGTTTACGGTTGGGGAGGGCGCACG 1299
          ||||| ||||||||                    |||||  |||||||||||||||||||
Db    187 CGGGGCGTCTTCGCGAGGCGAAAGTACGTTGGAGGAGATCTACGCGCCCGGGAGGGCCTACG 246

Qy   1300 GGTTGCTCATGGGGGTGGTGGGGAAGTTGCTCGGCGCACGTGATTCAGATTCTGGTGA 1359
          |||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Db    247 GGCTGTTCATGGGCGCGGCGGGGAAGCTCCTCGCCGCAGATCATCCAGATCCTGGTGA 306

Qy   1360 TTGCTGGCTGGGTTAGTGCGACCATGGGACCCTTGTTTGTTTGGGGTTGAATAAACTGAAGC 1419
          |||||||||||||||||||   ||||||| |  |||||||||||||||||||||   ||||
Db    307 TCGCCCGGGTGGGTGAGCTGCACCATGGGCCCCGGCTTCTTCTCACGCGGCTCAAGAAGTGGGCC 366

Qy   1420 TGTTGAGGATTTCTTCAGAGGATGAGCTTGCGCGGGATGGACATGACTCGCCATGGAGGCT 1479
          ||||||||||||||| |||||||||||| ||||||||||||||||||||||||||| ||||
Db    367 TGCTCGCGATCCTCGGCGCCGAGATGTCCGGCGAGATGTCCGCGCATGGACCTGACCTGGACCGGCACCGCGGCT 426

Qy   1480 TTGCTTATGCTTATGAGGATGATGA 1504
          ||||||||||||| |||||||||||
Db    427 TCGCCTACGTCTACCACGACGAGGA 451
```

RESULT 10
ABL94034/c
ID   ABL94034 standard; cDNA; 445 BP.
XX
AC   ABL94034;
XX
DT   10-JUN-2002 (first entry)
XX
DE   Arabidopsis thaliana nucleic acid sequence Ref:2027799 SEQ ID NO:799.
XX
KW   Arabidopsis thaliana; insecticide; fungicide; plant; mapping; diagnosis;
KW   genetic modification; gene; ss.
XX
OS   Arabidopsis thaliana.
XX
PN   US2002023280-A1.
XX
PD   21-FEB-2002.
XX
PF   26-JAN-2001; 2001US-00770444.
XX
PR   27-JAN-2000; 2000US-0178502P.
XX
PA   (GORL/) GORLACH J.
PA   (ANY/)   AN Y.
PA   (HAMI/) HAMILTON C M.
PA   (PRIC/) PRICE J L.
PA   (RAIN/) RAINES J M.
PA   (YUYY/) YU Y.
PA   (RAME/) RAMEAKA J G.
PA   (PAGE/) PAGE A.
PA   (MATH/) MATHEW A V.
PA   (LEDF/) LEDFORD B L.
PA   (WOES/) WOESSNER J P.
PA   (HAAS/) HAAS W D.
PA   (GARC/) GARCIA C A.
PA   (KRIC/) KRICKER M.
PA   (SLAT/) SLATER T.
PA   (DAVI/) DAVIS K R.
PA   (ALLE/) ALLEN K.
PA   (HOFF/) HOFFMAN N.
PA   (HURB/) HURBAN P.
XX
PI   Gorlach J,   An Y,   Hamilton CM,   Price JL,   Raines TM,   Yu Y;
PI   Rameaka JG,   Page A,   Mathew AV,   Ledford BL,   Woessner JP,   Haas WD;
PI   Garcia CA,   Kricker M,   Slater T,   Davis KR,   Allen K,   Hoffman N;
PI   Hurban P;
XX
DR   WPI; 2002-267486/31.
XX
PT   New Arabidopsis thaliana nucleic acid, for identifying homologous genes,
PT   producing compositions that modulate the expression or function of its
PT   encoded protein, and mapping functional regions of a protein.
XX
PS   Claim 1; SEQ ID NO 799; 44pp; English.
XX
CC   The present invention describes an Arabidopsis thaliana nucleic acid (I)
CC   comprising a sequence capable of hybridising under stringent conditions
CC   to a sequence (S1) selected from any one of the 999 sequences given in
CC   ABL93236 to ABL94234. (I) have insecticide and fungicide activities, and
CC   they can be used as protein expression modulators. (I) can be used in
CC   identifying homologous or related genes, in producing compositions that
CC   modulate the expression or function of their encoded proteins, mapping
CC   functional regions of the proteins, and in studying associated
CC   physiological pathways. (I) can also be used: (1) for the genetic
CC   manipulation of cells, particularly plant cells; (2) in screening assays
CC   of various plant strains to determine the strains that are best capable
CC   of withstanding a particular disease or environmental stress; (3) for
CC   enhancing or inhibiting production of a biosynthetic product in a plant;
CC   (4) as probes in mapping and in diagnosis, in genetic modification and
CC   for screening purposes, to generate additional copies of the nucleic
CC   acids, to generate ribozymes or antisense oligonucleotides, and as single
CC   -stranded DNA probes or as triple-strand forming oligonucleotides; and (5)
CC   for generating genetically modified transgenic organisms. Note: The
CC   sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from the
CC   USPTO web site
XX
SQ   Sequence 445 BP; 156 A; 101 C; 73 G; 115 T; 0 U; 0 Other;

Query Match   7.1%; Score 121; DB 6; Length 445;
Best Local Similarity   64.4%; Pred. No. 2.4e-15;
Matches 181; Conservative   0; Mismatches 100; Indels   0; Gaps   0;

```
Qy   1291 GGGCGCACGGGTTGCTCATGGGGGTGGTGGGGAAGTTGCTGGCCGGCGCACGTGATTCAGA 1350
          |||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db    441 GCGTCCGCGGTTGTTGTTTATGGGGGTGGTGGAGGAAAACTACTTGGAGCTCAGCTGATTCAGA 382

Qy   1351 TTCTGGTGATTGCTGGCTGGGTGGGTGGTTAGTGCGACCATGGGACCCTTGTTTGGGGGTTGAATA 1410
          ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db    381 TCATTGTGATCACGGGTGGGTAAGTGCGACCATGGGACCACTTTCTTCATCCTCAAGA 322

Qy   1411 AACTGAAGCTGTTGAGGATTTCTTCAGAGGATGAGCTTGCGGGGATGGACATGACTCGCC 1470
          | |||  |||||| ||||||||||||||||||      ||||||| |||||||||||||
Db    321 AAATGAAATTGTGCGGATATCGTCCGAGGATGGAGGATGGATGGATATGACCAGGC 262

Qy   1471 ATGGAGGCTTTGCTTATGCTTATGAGGATGATGAGACGGACCGGACCATGGGATGCAGTTGA 1530
          || |||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    261 ACGGTGGTTTTGCTTATATGTACTTTGATGATGAGTCTCACAAAGCCATTCAGCTTA 202

Qy   1531 GGAGGGGTTGGGCCCAACGCGGTCTTCCACCACCACCACTGAT 1571
          ||| ||||||||||||||||||||||||||||||||||||||
Db    201 GGAGAGTTGAGCACGATCTCCTTCCTTCTGTGTGCTAAT 161
```

RESULT 11
ABQ44816/c
ID   ABQ44816 standard; DNA; 985 BP.
XX
AC   ABQ44816;
XX
DT   12-JUL-2002 (first entry)

```
XX
DE   Oligonucleotide for detecting cytosine methylation SEQ ID NO 31407.
XX
KW   Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
KW   drug; side effect; cancer; central nervous system; cardiovascular;
KW   gastrointestinal; respiratory system; single nucleotide polymorphism;
KW   SNP; cell differentiation; ds.
XX
OS   Homo sapiens.
XX
XX   WO200218632-A2.
PN
XX
PD   07-MAR-2002.
XX
PF   01-SEP-2001; 2001WO-EP010074.
XX
PR   01-SEP-2000; 2000DE-01043826.
PR   05-SEP-2000; 2000DE-01044543.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Olek A, Piepenbrock C, Berlin K, Guetig D;
XX
DR   WPI; 2002-371829/40.
XX
PT   Determining the degree of cytosine methylation in genomic DNA, useful for
PT   diagnosis and prognosis, comprises selective hybridization of amplicons
PT   from chemically treated DNA.
XX
PS   Claim 12; 56pp + Sequence Listing; 56pp; German.
XX
CC   This invention describes a novel method for determining the degree of
CC   methylation of a particular cytosine in a motif 5'-CpG-3', present in a
CC   genomic sample of DNA. The sample is treated chemically to convert
CC   cytosine (C) but not methylated C, to uracil, then part of the genomic
CC   DNA that contains the target C is amplified to form a labeled amplicon.
CC   The amplicon is hybridised to two classes, each with at least one member,
CC   of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the
CC   degree of hybridisation to both classes is determined from the label on
CC   the amplicon. From the ratio of labels hybridised to the two classes of
CC   oligomers, the degree of methylation is calculated. The method is used:
CC   (i) for diagnosis and/or prognosis of side effects of therapeutic drugs
CC   and of a wide range of diseases, e.g. cancer, disorders of the central
CC   nervous, cardiovascular, gastrointestinal and respiratory systems etc.,
CC   particularly by detecting mutations or single nucleotide polymorphisms
CC   (SNP's); and (ii) for differentiation of cell or tissue types and for
CC   investigating cell differentiation. The method allows the methylation
CC   status of many C residues to be determined simultaneously. ABQ13410-
CC   ABQ54121 represent genomic DNA sequences used to illustrate the method
CC   for determining the degree of cytosine methylation described in the
CC   disclosure of the invention
XX
SQ   Sequence 985 BP; 234 A; 145 C; 326 G; 280 T; 0 U; 0 Other;

Query Match     5.2%;  Score 89;  DB 6;  Length 985;
Best Local Similarity  49.7%;  Pred. No. 1.1e-08;
Matches  227;  Conservative  0;  Mismatches 230;  Indels  0;  Gaps  0;

Qy    136 ACGCCTCCGCGCGCCGCCTCCCTTATCTGGGCCATTTCGCCGCCGTGGACAGCAAGTTCG 195
              ||||||| ||||||| |||||||| ||  |||||||| ||||||||||| |||||||
Db    600 ACATCGACGTCGACATCTACGTCGACGTCAACGTCGACATCGACGTCGACATCTACGTCG 541

Qy    196 TCGACACGGCCTTCGCCGTCGACAACACCTACCTCCTCTTTTCCGGCCTACCTCGTTTTTT 255
              |||||||| |||||||| |||||||| |||||||||||| |||||||| ||||
Db    540 ACGTCTACATCTACGTCGACGTCGACGTCGACGTCTACATCTACGTCTCACATCAACATCT 481

Qy    256 CTATGCAGCTCGGCCTTCGCCATGCTCGCGCCGGCGTCCGGTCCGGCCGGCCAAGAACACCATGA 315
              |||||| |||||||||||||||  |||| ||||| ||||||||||||||
Db    480 ACATCTACATCTACGTCGACGTCTACATCTACGTCGACGTCTCACATCTACGTGCT 421

Qy    316 ACATCATGTCACCAAGTCCTGGACGGCTGCGCCGCGGGCGCCCTTCTACTACCTCTTCG 375
              ||||||| |||||||||||||||| ||||||||||
Db    420 ACATCTACGTCTACATCTACGTCGACGTCGACGTCGACATCTCGACATCTCGACATCT 361
```

```
         376 GCTTCGCCTTGCTTTCGGCTCCCCTCCCAACGGCTTCATCGGTAAACATTTCTTCGGCC 435
Qy           |||||||| || ||| |||||||| |||||| |||||||||||||||||||||
Db    360 ACGTCGACATCGACATCGACATCGACGTCGACATCTACGTCAACGTCTCACATCT 301

         436 TCAAGGACATCCCTACTCATCCTCCGACTACAGCTACTTCCTCTACCAATGGCCTTCG 495
Qy           ||||||   |||||||
Db    300 ACGTCTACATCTACGTCGACATCTACGTCAACGTCGACATCGACATCTACGTCAACGTCG 241

         496 CCATGCGCCGCGCGGCGATCACCAGCGGAAGCATGCCGAACGCACAGTTCGTGGCCT 555
Qy           |||||||||||||  |||||||
Db    240 ACATCGACATCAACGTCGACATCAACATCTACGTCGACATCTACGTCGACATCT 181

         556 ATCTCATCTACTCCTCCTTCCTCACCGGCTTCGTCTA 592
Qy           ||||||||||
Db    180 ACGTCTCATCAACATCAACACATCTACGTCGACGTCTA 144
```

```
RESULT 12
ABQ44817
ID   ABQ44817 standard; DNA; 985 BP.
XX
AC   ABQ44817;
XX
DT   12-JUL-2002  (first entry)
XX
DE   Oligonucleotide for detecting cytosine methylation SEQ ID NO 31408.
XX
KW   Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
KW   drug; side effect; cancer; central nervous system; cardiovascular;
KW   gastrointestinal; respiratory system; single nucleotide polymorphism;
KW   SNP; cell differentiation; ds.
XX
OS   Homo sapiens.
XX
XX   WO200218632-A2.
PN
XX
PD   07-MAR-2002.
XX
PF   01-SEP-2001; 2001WO-EP010074.
XX
PR   01-SEP-2000; 2000DE-01043826.
PR   05-SEP-2000; 2000DE-01044543.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Olek A, Piepenbrock C, Berlin K, Guetig D;
XX
DR   WPI; 2002-371829/40.
XX
PT   Determining the degree of cytosine methylation in genomic DNA, useful for
PT   diagnosis and prognosis, comprises selective hybridization of amplicons
PT   from chemically treated DNA.
XX
PS   Claim 12; 56pp + Sequence Listing; 56pp; German.
XX
CC   This invention describes a novel method for determining the degree of
CC   methylation of a particular cytosine in a motif 5'-CpG-3', present in a
CC   genomic sample of DNA. The sample is treated chemically to convert
CC   cytosine (C) but not methylated C, to uracil, then part of the genomic
CC   DNA that contains the target C is amplified to form a labeled amplicon.
CC   The amplicon is hybridised to two classes, each with at least one member,
CC   of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the
CC   degree of hybridisation to both classes is determined from the label on
CC   the amplicon. From the ratio of labels hybridised to the two classes of
CC   oligomers, the degree of methylation is calculated. The method is used:
CC   (i) for diagnosis and/or prognosis of side effects of therapeutic drugs
CC   and of a wide range of diseases, e.g. cancer, disorders of the central
CC   nervous, cardiovascular, gastrointestinal and respiratory systems etc.,
CC   particularly by detecting mutations or single nucleotide polymorphisms
CC   (SNP's); and (ii) for differentiation of cell or tissue types and for
CC   investigating cell differentiation. The method allows the methylation
CC   status of many C residues to be determined simultaneously. ABQ13410-
```

CC   ABQ54121 represent genomic DNA sequences used to illustrate the method
CC   for determining the degree of cytosine methylation described in the
CC   disclosure of the invention
XX
SQ   Sequence 985 BP; 280 A; 326 C; 145 G; 234 T; 0 U; 0 Other;

Query Match   5.2%; Score 89; DB 6; Length 985;
Best Local Similarity   49.7%; Pred. No. 1.1e-08;
Matches 227; Conservative   0; Mismatches 230; Indels   0; Gaps   0;

Qy   136 ACGGCTCCGCCGCCGGCCTCCCTTATCTGCGGCCATTCGCCGCCGTGGACAGCAAGTTCG 195
                                                      ||||| |||
Db   386 ACATGACGTCGACATCTACGTCAACGTCAACGTCGACATCGACATCTACGTCG 445

Qy   196 TCGGACACGGCCTTGCCGTCGACAACACCTACCTCCTTTTTCGGCCTACCTCGTTTTTT 255
                                                      ||||| |||
Db   446 ACGTCTACATCTACGTCGACGTCGACGTCTACATCTACGTCTACATCAACATCT 505

Qy   256 CTATGCAGCTCGGCTTCGCCATGCTTCGCGCCGGCTCCGTCGCGCCAAGAACACCATGA 315
                                                      ||||| |||
Db   506 ACATCTACATCTACGTCGACGTCTACATCTACGTCCGACATCTACGTCTACATCTACGTCT 565

Qy   316 ACATCATGCTCACCAAGGTCCTGGACGCTGCCGGCGGCCTCTTCTACCTCTTCG 375
                                                      || |||
Db   566 ACATCTACGTCTACATCTACGTCACGTCGACGTCGACATCTACGTCGACATCT 625

Qy   376 GCTTCGGCTTCGCTTCGGTCTCCCCCTCCAACGGCTTCATCGTAAACATTTCTTCGGCC 435
                                                      ||||| |||
Db   626 ACGTCGACATCGACATCGACATCGACGTCTACATCTACGTCAACGTCTACATCT 685

Qy   436 TCAAGGACATCCCTTCATCCTCTACGACTACAGCTACTTCCTCTACCAATGGGCCTTCG 495
                                                      || ||
Db   686 ACGTCTACATCTACGTCGACATCTCAACGTCAACGTCGACATCTACGTCAACGTCG 745

Qy   496 CCATGCGCGCCGGCGGAGCATCCCGGAAGCATCGCCGAACGCACACAGTTCGTGGCCT 555
                                                      |||||
Db   746 ACATCGACATCAACGTGCGACATCGACATCAACAACATCTACGTCGACATCTACGTCG 805

Qy   556 ATCTCATCTACTCCTCCTTCCTCACCGGCTTCGTCTA 592
                                                      ||
Db   806 ACGTCTACATCAACATCAACATCTACGTCCACGTCTA 842

RESULT 13
ABL11211
ID   ABL11211 standard; cDNA; 1956 BP.
XX
AC   ABL11211;
XX
DT   26-MAR-2002 (first entry)
XX
DE   Drosophila melanogaster expressed polynucleotide SEQ ID NO 28115.
XX
KW   Drosophila; developmental biology; cell signalling; insecticide;
KW   pharmaceutical; gene; ss.
XX
OS   Drosophila melanogaster.
XX
PN   WO200171042-A2.
XX
PD   27-SEP-2001.
XX
PF   23-MAR-2001; 2001WO-US009231.
XX
PR   23-MAR-2000; 2000US-0191637P.
PR   11-JUL-2000; 2000US-0611450.
XX
PA   (PEXE) PE CORP NY.
XX
PI   Venter JC, Adams M, Li PWD, Myers EW;
XX
PR   WPI; 2001-656860/75.
PR   P-PSDB; ABB67108.

XX
PT   New isolated nucleic acid detection reagent for detecting 1000 or more
PT   genes from Drosophila and for elucidating cell signaling and cell-cell
PT   interactions.
XX
PS   Claim 1; SEQ ID NO 28115; 21pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated nucleic acid detection reagent
CC   capable of detecting 1000 or more genes from Drosophila. The invention is
CC   useful in developmental biology and in elucidating cell signalling and
CC   cell-cell interactions in higher eukaryotes for the development of
CC   insecticides, therapeutics and pharmaceutical drugs. The invention
CC   discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC   sequences (ABL01840-ABL16175) and the encoded proteins (ABB5737-
CC   ABB72072). The sequence data for this patent did not form part of the
CC   printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 1956 BP; 395 A; 536 C; 549 G; 476 T; 0 U; 0 Other;

Query Match   5.1%; Score 87.6; DB 4; Length 1956;
Best Local Similarity   55.6%; Pred. No. 2.4e-08;
Matches 218; Conservative   0; Mismatches 159; Indels   15; Gaps   2;

Qy   206 CTTCGCCGTCGACAACACCTACCTCCTTTTCCGCTACCTCGTTTTTCTATGCAGCT 265
                                                      |||||
Db   336 CTCCGTCGAGGACACCAACTGGGTCCTGACCTCCTTCATCATCTTCACCATGCAAAC 395

Qy   266 CGGCTTCGCCATGCTCTGCGCGGCTCCGTCGCGCCAAGAACACCATGAACATCATGCT 325
                                                      ||||||
Db   396 GGGTTTCGGTATGCTCGGAGTCCGGTTGTGTAAGCATCAAGAACGAGGTGAACATCATGAT 455

Qy   326 CACCAACGTCCTGACGCTGCCGCGCGGCGCCTCTTCTACTACCTCTTCGGCCTTCG---- 381
                                                      ||||||
Db   456 GAAGAACGTGATCGACATTGTCTGGGCGGATTACTTACTGGCTGCTCTTCGGCTACGGAAT 515

Qy   382 --CCTTCGCTTTCGGCTCCCCCTCCAACGGCTTCATCGGTAAACATTTCTTCGGCCTCAA 439
                                                      |||||
Db   516 GAGCTTCGGTGGTCGCGGGCCCCACTCTTCCAATCATTCATCGCCATTGGGGACTTCCTGCTGGA 575

Qy   440 --------GGACATCCCTTCATCCTCTACGACTACAGCTACTTTGCCGCCTTCCTGTTCCAACTGTC 490
                                                      |||||
Db   576 TCCACGGGTGGGCGATGCACTGATGGACAGATTTTTGCCGCCTTCCTTCCAACTGTC 635

Qy   491 CTTCGCCATCGCCGCGCCGCCGCATCCACCAGCGGAAGCATCGCCGAACGCACACAGTTCGT 550
                                                      ||||||
Db   636 GTTCGCCACCACGGCCACAACCATTGTCAGCGGCCGCCATGGTGAACGCTGCAACTTTAA 695

Qy   551 GGCCTATCTCATCTACTCCTCCTTCCTCACCG 582
                                                      ||||||
Db   696 GGCATACTGCCTGTTCTCCTTCCTGAACACCG 727

RESULT 14
AAD07130/c
ID   AAD07130 standard; cDNA; 2805 BP.
XX
AC   AAD07130;
XX
DT   06-AUG-2001 (first entry)
XX
DE   Canine retinitis pigmentosa GTPase regulator (RPGR) cDNA.
XX
KW   Dog; X-linked progressive retinal atrophy; XLPRA; genetic marker;
KW   retinitis pigmentosa GTPase regulator; RPGR; Siberian Husky; Samoyed;
KW   Miniature Schnauzer; ss.
XX
OS   Canis familiaris.
XX
FH   Key   Location/Qualifiers
FT   CDS   1..1569
FT   /*tag= a
FT   /product= "Canine retinitis pigmentosa GTPase regulator"

FT   /note= "CDS does not include start codon"
FT   /partial
XX
PN   WO200138578-A1.
XX
PD   31-MAY-2001.
XX
PF   21-NOV-2000; 2000WO-US031940.
XX
PR   24-NOV-1999;   99US-0167365P.
XX
PA   (CORR) CORNELL RES FOUND INC.
XX
PI   Aguirre GD, Acland GM, Zhang Q, Ray K, Zeiss CJ;
XX
DR   WPI; 2001-367707/38.
DR   P-PSDB; AAE02397.
XX
PT   Identifying dogs with or carrying X-linked progressive retinal atrophy by
PT   detecting retinitis pigmentosa GTPase regulator gene mutation, useful
PT   when breeding Husky, Samoyed and Miniature Schnauzer.
XX
PS   Claim 71; Page 30-31; 88pp; English.
XX
CC   The invention relates to a method for identifying dogs which are
CC   genetically normal, are carriers of, or are affected with X-linked
CC   progressive retinal atrophy (XLPRA), by testing a biological sample with
CC   genetic markers that co-segregate with a XLPRA gene locus. The invention
CC   also relates to canine retinitis pigmentosa GTPase regulator (RPGR), RPGR
CC   mutants and their corresponding nucleic acid molecules. The mutated RPGR
CC   genes are responsible for the XLPRA in dogs. Methods are used to select
CC   dogs for breeding so that dogs carrying the mutated locus are eliminated
CC   from the breeding stock. The method particularly applies to Siberian
CC   Husky, Samoyed and Miniature Schnauzer breeds, or any other breed where
CC   the disease is X-linked. XLPRA1 type is identified in Siberian Huskies,
CC   and Samoyeds, while XLPRA2 type is identified in Miniature Schnauzers.
CC   The present cDNA sequence is the normal open reading frame (ORF) 15
CC   encoding Canine retinitis pigmentosa GTPase regulator (RPGR) protein
CC   found in normal dogs
XX
SQ   Sequence 2805 BP; 1029 A; 287 C; 940 G; 549 T; 0 U; 0 Other;

Query Match   5.0%;   Score 85.2; DB 4; Length 2805;
Best Local Similarity   46.8%;   Pred. No. 8.3e-08;
Matches 267; Conservative   0; Mismatches 303; Indels   0; Gaps   0;

Qy   137 CGCCTCCGCCGCCGCCTCCTTATCTGCGGCCATTTCGCGCCGCCTGGACAGCAAGTTCGT 196
            | ||| |||||||| | | | | || ||||||||||||||||||||||||
Db   1123 CCCCCTCCTCCTCCTCCTCCTCCCCCCTCCTCCTCCCCCCTTCTTCCTCCTCCTCCTTCTT 1064

Qy   197 CGACAGGGCCTTCGCCGGTGACAAACACCTACCTCCTCTTTTCGGCCTACCTCGTTTTTC 256
            | || | ||||||| | |||| | | | |||| |||| |||||||
Db   1063 CTTCCCTCCTCCTCGGCCCTTCCCTCCTCCTCCTCCTCTCCCTTCTCCCCCTCCCCCCTT 1004

Qy   257 TATGCAGCTCGGCTTCGGCCATGCTCGGCCGGCTCGGTCCGGCCGGCCAAGAACACCATGAA 316
            | || ||||| | ||||| | | | | || | | ||||| |||||
Db   1003 CTTCCTGCCCTGCTTCCCCCTCTTCCTCCCCCTCCCCCTCCCTGCTCGTCTCCCCCTTCTT 944

Qy   317 CATCATGTCACCAACGTCCTGGACGCTGCGCGCGGCGGCCTCTTCTACTACCTCTTCGG 376
            | |||||| || ||| | | | ||| | | ||| | |||||
Db   943 CTTCCCTCCTCCTTCATTTCTCTTCTTTTCCTTTCCTCCTCCTCCTCCTTCCTCCCCCTTCAT 884

Qy   377 CTTCGCCTTCGGCTTCGGCTCCCCCTCCAACGGCTTCATCGGTAAACATTTCTTCGGCCT 436
            ||| | ||| | ||| ||||| | | | | | | | | ||||
Db   883 ATTCTCTTTTGTTTCCTTCCCCTCCCTCCTTCTTCCCACTTCTTCTTCTTCTTCTTCTG 824

Qy   437 CAAGGACATCCTTCATCCTCCTAGGACTACAGTAGTCCTCCAATGGCCTTCGC 496
            | | ||||| | ||||| | ||| | ||| || | ||||
Db   823 CATCCTCTTCTTCGCTTCTCTCCACTTCCCCCTCTCCCTTCTCCCTTCCTTTCTT 764

Qy   497 CATCGCGCGCCGGCATCACCAGCGGAAGCATGCCGAACGCACACAGTTCGTGGCCTA 556
            | ||| | | | | | | | ||||| | | | | | | | | |
Db   763 CCACTTCCTCCTTCCTCCCACTTCCCACTTCCTCCTCTCTCCACTTCCCCCTTTC 704

RESULT 15
ABL11210/c
ID   ABL11210 standard; cDNA; 5296 BP.
XX
AC   ABL11210;
XX
DT   26-MAR-2002 (first entry)
XX
DE   Drosophila melanogaster expressed polynucleotide SEQ ID NO 28112.
XX
KW   Drosophila; developmental biology; cell signalling; insecticide;
KW   pharmaceutical; gene; ss.
XX
OS   Drosophila melanogaster.
XX
PN   WO200171042-A2.
XX
PD   27-SEP-2001.
XX
PF   23-MAR-2001; 2001WO-US009231.
XX
PR   23-MAR-2000; 2000US-0191637P.
PR   11-JUL-2000; 2000US-00614150.
XX
PA   (PEKE) PE CORP NY.
XX
PI   Venter JC, Adams M, Li PWD, Myers EW;
XX
DR   WPI; 2001-656860/75.
DR   P-PSDB; ABB7107.
XX
PT   New isolated nucleic acid detection reagent for detecting 1000 or more
PT   genes from Drosophila and for elucidating cell signaling and cell-cell
PT   interactions.
XX
PS   Claim 1; SEQ ID NO 28112; 21pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated nucleic acid detection reagent
CC   capable of detecting 1000 or more genes from Drosophila. The invention is
CC   useful in developmental biology and in elucidating cell signalling and
CC   cell-cell interactions in higher eukaryotes for the development of
CC   insecticides, therapeutics and pharmaceutical drugs. The invention
CC   discloses genomic DNA sequences (ABL1840-ABL16175), expressed DNA
CC   sequences (ABL01840-ABL16175) and the encoded proteins (ABB5737-
CC   ABB72072). The sequence data for this patent did not form part of the
CC   printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 5296 BP; 1425 A; 1165 C; 1199 G; 1507 T; 0 U; 0 Other;

Query Match   4.6%;   Score 78.6; DB 4; Length 5296;
Best Local Similarity   54.9%;   Pred. No. 2.2e-06;
Matches 206; Conservative   0; Mismatches 154; Indels   15; Gaps   2;

Qy   206 CTTCGCGTCGACAACACCTACCTCCTCTTTTCCGCCTACCTCGTTTTTCTATGCAGCT 265
            ||| |||| | ||| ||| ||| |||| ||||||||
Db   2823 CTTCCGTCGAGGACGACCAACTGGGTCCTGACCTCCTCATCATCATCTTCACCATGCAAAC 2764

Qy   266 CGGCTTCGGCCATGCTCGCGCCGGCTCCGGCCGGGCCAAGAACACCATGACACATCATGCT 325

Qy   557 TCTCATCTACTCCTTCCTTCCACCGGCTTCGTCATCCGGTGGTGGTCCCCACTGGTTCTG 616
            ||||||||| | | ||| ||| |||||||
Db   703 CTTCCCATTCTGCCTTCCTTCCTCCCTCCTCCAGCCTCCTTCTCTCCTCCCCTTCCTCTT 644
XX
Qy   617 GTCCCCAGACGGCTGGGCCTTCTGCCTTAAGATCACCGACCGGCTATTTCCACCGGCGT 676
            ||| | || | || | | ||| | | || ||| || | ||| ||||||
Db   643 CCCCTCCTCCTCCTCCCTCCAATTCCCCCTCTTCGTCTTCTCCCTTTCCCTCTCCT 584
XX
Qy   677 AATAGACTTCGCCGGTTCGCGGGTAGTCCA 706
            | || || | ||| | |||||
Db   583 CCCTAATTCCCCCTCCCCCTTCCTCTTCCA 554

```
Db   2763 GGGTTTCGGTATGCTGGAGTCCGGTTGTGTAAGCATCAAGAACGAGGTGAACATCATGAT 2704
            ||| ||||| |||||          |||| |  ||       |  ||||||||||||||
Qy    326 CACCAACGTCCTGGACGCTGCCGCCGGCGCCTCTTCTACTACCTCTTCGGCTTCG----  381
            |  ||||| ||||  ||         |||||| |    ||| |||||||||||
Db   2703 GAAGAACGTGATCGACATTGTTCTGGGCGGATTTACTTACTGGCTCTTCCGGCTACGGAAT 2644
                                        ||||||          ||||||||||||||
Qy    382 --CCTTCGCTTTCGGCTCCCCTCCAACGGCTTCATCGGTAAACATTTCTTCGGCCTCAA  439
            |||||||           ||||||||| |||  |||||||
Db   2643 GAGCTTCGGTCGCGGCCCACTCTCCAATCCATTCATCGGCCATTGGGGACTTCCTGCTGGA 2584
                                                   ||||||  ||
Qy    440 --------GGACATCCCTTCATCCTCCTACGACTACAGCTACTTCCTCTACCAATGGGGC  490
                      |                ||||| ||||| |  ||||| |
Db   2583 TCCACCGGTGGGCGATGCCACTGATGGGACAGATTTTGCCGCCTTCCTGTTCCAACTGTC 2524
                                                     ||||||  |   |||||
Qy    491 CTTCGGCCATCGCCGCGCCGGCATCACCAGCGGAAGCATCGCCGAACGCACACAGTTCGT  550
            ||||||||  |  ||||         ||||||       |||||
Db   2523 GTTCGCCACCACGGCCACAACCATTGTCAGCGGCGCCCATCGCTGAACGGTGAGTAACCAT 2464
            ||||||||||  ||||||           ||
Qy    551 GGCCTATCTCATCTA  565
            |||||||||| |||
Db   2463 GATACATCTCTACTA 2449

Search completed: March  9, 2004, 10:48:21
Job time : 758 secs
```

```
                    GenCore version 5.1.6
            Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:        March 9, 2004, 10:35:31 ; Search time 155 Seconds
                                          (without alignments)
                                          6108.040 Million cell updates/sec

Title:         US-10-033-109-3
Perfect score: 1706
Sequence:      1 gcacgagcactcccaaccc........aaaaaaaaaaaaaaaa 1706

Scoring table: IDENTITY_NUC
               Gapop 10.0 , Gapext 1.0

Searched:      682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters:      1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :    Issued_Patents NA:*
           1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
           2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
           3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
           4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
           5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
           6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                          SUMMARIES
```

Result No.	Score	% Query Match	Length	DB	ID	Description
1	755.8	44.3	1748	4	US-08-635-967-1	Sequence 1, Appl
2	91.4	5.4	1926	4	US-09-249-585A-4	Sequence 4, Appli
3	91.4	5.4	1931	2	US-09-130-114-2	Sequence 1, Appl
4	82	4.8	7218	1	US-08-232-463-14	Sequence 14, Appl
5	76.6	4.5	390	3	US-09-197-649-7	Sequence 7, Appl
6	76	4.5	1926	4	US-09-249-585A-2	Sequence 2, Appli
7	76	4.5	1926	4	US-09-410-399-3	Sequence 3, Appli
8	76	4.5	2580	4	US-09-359-863-2	Sequence 2, Appli
9	76	4.5	2580	4	US-09-359-081-2	Sequence 2, Appli
10	76	4.5	5452	2	US-09-130-114-1	Sequence 1, Appl
11	76	4.5	8705	3	US-09-647-344A-14	Sequence 14, Appl
12	76	4.5	9600	3	US-08-910-641-1	Sequence 1, Appl
13	76	4.5	9600	4	US-08-620-925-1	Sequence 1, Appl
14	76	4.5	10596	1	US-07-884-811-15	Sequence 15, Appl
15	76	4.5	10596	1	US-07-885-971-15	Sequence 15, Appl
16	76	4.5	10596	1	US-08-087-783A-15	Sequence 15, Appl
17	76	4.5	10596	2	US-08-194-088B-15	Sequence 15, Appl
18	76	4.5	10596	1	US-08-194-087-15	Sequence 15, Appl
19	76	4.5	10596	5	PCT-US93-04648-15	Sequence 15, Appl
20	76	4.5	16080	4	US-09-724-566A-48	Sequence 48, Appl
21	74.4	4.4	4403765	3	US-09-103-840A-2	Sequence 2, Appli
22	74.4	4.4	4411529	3	US-09-103-840A-1	Sequence 1, Appli
23	74.2	4.3	1248	3	US-09-105-537-7	Sequence 7, Appl
24	74.2	4.3	13613	3	US-09-105-537-3	Sequence 3, Appli
25	68	4.0	1505	1	US-07-915-246-1	Sequence 1, Appl
26	66.6	3.9	5970	4	US-09-320-878-21	Sequence 21, Appl
27	66.6	3.9	5970	4	US-09-141-908-11	Sequence 11, Appl

c 28		3.9	5970	4	US-09-657-440-21	Sequence 21, Appl
29	66.6	3.8	1965	3	US-09-178-252-26	Sequence 26, Appl
30	65.4	3.8	1965	4	US-09-826-660-26	Sequence 26, Appl
31	65.4	3.8	16442	4	US-08-781-891-208	Sequence 208, App
32	64.8	3.8	16442	4	US-09-618-166-208	Sequence 208, App
c 33	64.8	3.8	4403765	3	US-09-103-840A-2	Sequence 2, Appli
c 34	64.4	3.8	4411529	3	US-09-103-840A-1	Sequence 1, Appli
c 35	64.4	3.7	3489	4	US-08-728-323A-1	Sequence 1, Appli
c 36	63	3.7	3489	4	US-09-298-568-1	Sequence 1, Appl
c 37	63	3.7	3489	4	US-09-410-399-1	Sequence 1, Appl
38	63	3.7	32207	2	US-08-770-379-20	Sequence 20, Appl
39	63	3.7	32207	3	US-08-757-669A-20	Sequence 20, Appl
40	63	3.7	32207	4	US-09-230-371A-20	Sequence 20, Appl
41	62.8	3.7	4466	4	US-09-410-551B-16	Sequence 16, Appl
42	62.8	3.7	4478	4	US-09-410-551B-20	Sequence 20, Appl
43	62.8	3.7	4547	4	US-09-410-551B-22	Sequence 22, Appl
44	62.8	3.7	4571	4	US-09-410-551B-18	Sequence 18, Appl
45	62.8	3.7	77536	4	US-09-410-551B-1	Sequence 1, Appli

```
                          ALIGNMENTS

RESULT 1
US-08-635-967-1
; Sequence 1, Application US/08635967
; Patent No. 6620610
; GENERAL INFORMATION:
;;  APPLICANT: FROMMER, Wolf-Bernd
;;  APPLICANT: NINNEMAN, Olaf
;;  TITLE OF INVENTION:  DNA SEQUENCES FOR AMMONIUM TRANSPORTER,
;;  TITLE OF INVENTION:  PLASMIDS, BACTERIA, YEASTS, PLANT CELLS AND PLANTS
;;  TITLE OF INVENTION:  CONTAINING THE TRANSPORTER
;;  NUMBER OF SEQUENCES:  2
;;  CORRESPONDENCE ADDRESS:
;;   ADDRESSEE: Ostrolenk, Faber, Gerb & Soffen, LLP
;;   STREET: 1180 Avenue of the Americas
;;   CITY: New York
;;   STATE: NY
;;   COUNTRY: US
;;   ZIP: 10036-8403
;;  COMPUTER READABLE FORM:
;;   MEDIUM TYPE: Floppy disk
;;   COMPUTER: IBM PC compatible
;;   OPERATING SYSTEM: PC-DOS/MS-DOS
;;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;;  CURRENT APPLICATION DATA:
;;   APPLICATION NUMBER: US/08/635,967
;;   FILING DATE: 29-APR-1996
;;   CLASSIFICATION: 800
;;  PRIOR APPLICATION DATA:
;;   APPLICATION NUMBER: PCT/EP 94/03499
;;   FILING DATE: 24-OCT-1994
;;  PRIOR APPLICATION DATA:
;;   APPLICATION NUMBER: DE P 43 37 597.9
;;   FILING DATE: 28-OCT-1993
;;  ATTORNEY/AGENT INFORMATION:
;;   NAME: Meilman, Edward
;;   REGISTRATION NUMBER: 24,735
;;   REFERENCE/DOCKET NUMBER: P/951-120
;;  TELECOMMUNICATION INFORMATION:
;;   TELEPHONE:  (212) 382-0700
;;   TELEFAX:  (212) 382-0888
;;   TELEX: 236925
;;  INFORMATION FOR SEQ ID NO: 1:
;;   SEQUENCE CHARACTERISTICS:
;;    LENGTH: 1748 base pairs
;;    TYPE: nucleic acid
;;    STRANDEDNESS: single
;;    TOPOLOGY: linear
;;   MOLECULE TYPE: cDNA
;;   ORIGINAL SOURCE:
;;    ORGANISM: Arabidopsis thaliana
```

INDIVIDUAL ISOLATE: Ammonium transporter
IMEDIATE SOURCE:
  LIBRARY: cDNA library in plasmid pF161
POSITION IN GENOME:
  MAP POSITION: from 21 to 1523 coding region
FEATURE:
  NAME/KEY: CDS
  LOCATION: 21..1526
S-08-635-967-1

Query Match   44.3%;   Score 755.8;   DB 4;   Length 1748;
Best Local Similarity 71.7%;   Pred. No. 1.7e-173;
Matches 1036; Conservative   0; Mismatches 397;   Indels   12;   Gaps   3;

```
Qy  136 ACGCCTCCGCCGCCGCCTCCCTTATCTGCGGCCATTCGCGCCGCCGTGGACAGCAAGTTCG 195
            ||||||||||||  |||||   |||||| |||  ||||||||| ||| | |||||||||
Db   64 ATGCCACGCGCGGCCAACTACATAGTGCCAGCTAGGCGACGTCAACAACAAATTCA 123

Qy  196 TCGACACGGCCTTCGCGTCGACAACACTCACCTCCTCTTTCCGCCTACCTCGTTTTT 255
            |||||||||||||||  ||||  || |||||| ||||||||||||||||||||
Db  124 TCGACACCCGCTTTCGCTATAGACAACACTTACCTCCTCTTCTCCGCCTACCTTGTCTTCT 183

Qy  256 CTATGCAGCTCGGCTTCGCCATGCTTCGCGCGGCGTCCGTCCGCCAAGAACACCATGA 315
            ||||||||||||||||||| ||||| |||| |||||| ||||||||| ||| |||||
Db  184 CTATGCAGCTTGGCTTCGCTATGCTTCGTGTGCCGGTTCCGTGAGAGCCAAGAATACTATGA 243

Qy  316 ACATCATGCTCACCAACGTCCTGGACGCTGCCGGCCGCCGGCCTCTTCTACTACCTCTTCG 375
            |||||||||| |||||||||||||| |||| |||||||| |||||| || ||||||||
Db  244 ACATCATGCTTACCAACGTCCTTGACGCTGGACGCGCCGGTGGTCTCTTCTATATCTGTTTG 303

Qy  376 GCTTCGCCTTCGCTTTCGGGTTCCCCCCTCCAACGGCTTCATCGGTAAACATTTCTTCGGCC 435
            ||||||||||||||||| |||||||| ||||||||||| |||||| |||||||| |||||
Db  304 GCTACGCCTTTGCCTTGGATCTCCGTCCAATGGTTTCATCGGTAAACACTACTTTGGTC 363

Qy  436 TCAAGGACATCCTTCATCCTCCTACGACTACAGCTACTTCCTCTACCAATGGCCTTCG 495
            |||| |||  ||||| ||||||| ||||||||||||| ||||||| || ||||||||
Db  364 TCAAAGACACATCCCACGGGCCTCTGCTGACTACTCCAACTTTCTCTACCAATGGGCCTTTG 423

Qy  496 CCATCGCCGCCGCGGCATCACCAGCGGAAGCATGCCGAACGCACACAGTTCGTGGCCT 555
            |||||||||||| ||||||||||| ||||| ||||||||| ||||||||||||||||
Db  424 CAATCGCTCGCGGCTGGAATCACAAGGGCCTGCGATCGCTGAACGGACACACAGTTCGTGGCTT 483

Qy  556 ATCTCATCTACTCCTTCCTCACCGGCTTCGTCTATCCGGTGGTCTCCCACTGGTTCT 615
            ||||||||||||||||||| ||||||||||| |||||||||||||||||||||||||
Db  484 ACCTAATCTATTCCTTCTTCTTAACCGGGTTGTTTACCCGGCGTCGTCTCTCACTGGTTCT 543

Qy  616 GGTCCCCAGACGCGGCGGGCCTTCGCTGCTT---TAAGATCACCGACCGGCTATTTTCCACCG 672
            ||||      |||||||   ||||| ||||     ||| ||||| ||||| |||||||||
Db  544 GGTCAGTTGATGGATGGGCCAGCCCGTTCCGTACCGATGGAGATTTGCTTTTCAGCACCG 603

Qy  673 GGTAATAGACTTCGCCGGTTCCGGCGTAGTCCACATGGTCGGCGAATAGCCGGCCTAT 732
            |||| ||||||||||||| |||||  |||||||||||| |||||| ||||||||||||
Db  604 GAGCCGATAGATTTCGCTGGGTCCGGTGTTGTTCATATGGTCGGAGGTATCGCTGGACTCT 663

Qy  733 GGGGAGCGCGTGATCGAAGGCCAAGAATGGGACGTTTCGATCATGCAGGAGGACAGCTGTGG 792
            ||||  |||| |||||||| ||||| |||||| ||||||||||||| |||||||||||||
Db  664 GGGGTGCGCTCATCGAAGGTCCACGACTTCGGCCGGTTCGATAACGGAGGCCGTGCCATCG 723

Qy  793 CCTTGCGAGGCCACAGCGCGGTCCTTAGTAGTCCTGGAACCTTCTTGCTTGGTTTCGGTT 852
            || || |||| |||| |||| |||||||||  ||| || ||| |||||||| |||||
Db  724 CTCTTCGTGGCCACTCGGCGTCACTTGTTGTCCTTGGAACATTCCTCCTCTGGTTTTGGAT 783

Qy  853 GGTACGGATTTAACCCCGGTTCATTAACAAATCCTACTTACGGTAACTCAGGAA 912
            ||||||||||||||||||||||||| ||| ||||||| |||||||||| |||||
Db  784 GGTACGGATTTAACCCCGGTTCCTTCAACAAGATCCTAGTCAACGTACGAGA---CAGGCA 840

Qy  913 ATTACTACGGTCAATGGAGGCGCGGTTGGCAGAACCGCGGTCACCACTACCCTAGGCGGGGT 972
             ||||  |||||| ||||| | ||||| ||||||||| |||||| |||||||||||||||
Db  841 CATACAACGCGACGCCAGTGAGCGCGGTCGGACGGCAGTGTCACAACAACGTTAGCTGCT 900

Qy  973 CAACAGCTGCCTTGACCACGCTATTCGGTAGAGGGGTGATATCGGTGATATCCGGTCACTGGAACGTGA 1032
            |||||||||||||||||||||| ||||| |||| || | ||||||| |||||||||| |||||||||
Db  901 GCACCGCGGCGCTGACAACCCTAAGCCTACTCCGGGACATTGGAACGTCA 960
```

```
Qy 1033 CCGATGTCTGCAACGGCGGCTGTTAGGCGGTTCGCGGCGCGATAACAGCCGGTTGCTCCGTGG 1092
            |||| || | |||||||||||  ||    ||| | ||  ||    |||||||||||| |||
Db  961 CTGATGTATGCAACGGCCCTCCCTCCGGAGGGTTTGCAGCCATAACTGGTGGCTGCTTGTCG 1020
```

```
Qy 1093 TTGAGCCATGGCAGCCATCGTATGCGGTTTTGTTGTTCTTCTATAGTATTAATAGCTTGCA 1152
            |||||||| |||||||| | ||||||||| | ||||||| ||||| |||||||||||||
Db 1021 TTGACCATGGCTGCGATCATCTCGGGGTTCGTCGGGGGTCCTCAGTCCTCCTCGGATGCA 1080
```

```
Qy 1153 ACAAATTAGCAGAGAAGGTTAAGTTCGACGATCCTCTGGAGGCGGCGGCGAGTTGCACGGTG 1212
            |||  |  |||||||||||||||||| ||| ||| |||||||| || | |||||||||||
Db 1081 ACAAGCTCGCTGAGAAGCTCAAATACGACGACCCCTCTTGAGGCAGCACAACTACACGGTG 1140
```

```
Qy 1213 GGTGTGGCACGTGGGGGGTGATATTCACGGCCGTTGTTCGCAAAAAGGAGTATGTGAAGG 1272
            |||||||||||| |||||| |||  ||| ||||||||||||||  |||||||| ||||||
Db 1141 GTTGGGGTGGCGTGGGGGACTAATATTCACGGGCTCTCTCGCTCGTCAAGAAAAGTACTTGAACC 1200
```

```
Qy 1273 AGGTTTACGGGTTG-----GGGAGGGCGCACGGGTTGCTCATGGGGGGTGTGGGAAGT 1326
            ||||||||||||       |||||||| ||| ||||| | |||| |||| ||| |||
Db 1201 AGATTTACGGCAACAAACCCGGAGGCCACACGGTTTGTTTATGGGCGTGGGAGGAAAAC 1260
```

```
Qy 1327 TGCTGGCGGCGCACGTGATTCAGATTCGGTGATTGGTGGGTGGGTAGTGGGACCATGG 1386
            ||||||||||| | |||| ||| ||| ||||| ||| || |||| ||| |||||||||
Db 1261 TACTTGGAGCTCAGCTGATTCAGATCATTGTGATCACGGGTTGGGTAAGTGCGACCATGG 1320
```

```
Qy 1387 GACCCTTGTTTGTGGGGTTGAATAAACTGAAGCTGTTGAGGATTCTTCAGAGGATGAGC 1446
            ||||| || ||| |||| ||| | |||| ||  ||||| |||||||||| ||||||||
Db 1321 GGACACTTTTCTCATCCTCAAGAAATGAAATTGTGCGGATATCGTCCGAGGATGAGA 1380
```

```
Qy 1447 TTGCGGGGATGGACATGACTCGCCATGGAGGCTTTGCTTATGCTTATGAGGATGATGAGA 1506
            ||||||||| | | |||||| | |||  || || |||| |||| |||||||| ||||||
Db 1381 TGGCCGTATGGATATGACCAGCAGGCACGTGGTTTTGCTTATATGTACTTTGATGATGATG 1440
```

```
Qy 1507 CGCACAAGCATGGGATGCAGTTGAGGAGGTTGGGCCCAACCCGTCTTCCACACCCACCA 1566
            |||| || | |||  |||||| ||||| |||||||||  |||||||||||||||||||||
Db 1441 AGTCTCACAAAGCCATTCAGCTTAGGAGAGTTAGGAGAGTTGAGCCACGATCTCCTTCTCTGGTG 1500
```

```
Qy 1567 CTGAT 1571
            || ||
Db 1501 CTAAT 1505
```

RESULT 2
US-09-249-585A-4
; Sequence 4, Application US/09249585A
; Patent No. 6417002
; GENERAL INFORMATION:
; APPLICANT: Horlick, Robert
; TITLE OF INVENTION: METHOD FOR MAINTENANCE AND SELECTION OF EPISOMES
; FILE REFERENCE: 0867/0D905
; CURRENT APPLICATION NUMBER: US/09/249,585A
; CURRENT FILING DATE: 1999-02-11
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.0
;
; SEQ ID NO 4
; LENGTH: 1926
; TYPE: DNA
; ORGANISM: Epstein Barr Virus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1926)
; OTHER INFORMATION: template strand of EBNA-1 DNA
US-09-249-585A-4

Query Match   5.4%;   Score 91.4;   DB 4;   Length 1926;
Best Local Similarity 46.7%;   Pred. No. 1.4e-12;
Matches 290; Conservative   0; Mismatches 331;   Indels   0;   Gaps   0;

```
Qy  110 CCAACTTCTCGCCCAAACACCACCAGACGCCTCCGCCGCCGCCTCCCTTATCGCGGCCA 169
            || |||| ||||| ||| || ||   ||||  ||||||||||||||||||||| |||
Db  259 CCACCTTGTCCTCGTCCCCGCCTCCCCCGTCTCGCCCGTCTCCCCGTCTTCGCCGTCCT 318
```

```
Qy  170 TTTCGGCGCCCTGAGCAGCAGCAAGTTCGTCGACAACGCCCTTCGCCGTCGACAACACTACCT 229
            |||||||||||| || | ||| | ||||| |||| || || || ||| | || ||||
Db
```

```
Qy  230 CCTCTTTTCCGCCTACCTCGTTTTTCTATGCAGCTGGCTTCGCCATGCTCTGCGCCGG 289
            |||||  |  ||||       |           ||
Db  379 CCTCCTCCCGTCCTCGTCCTCCCCGTCCTCCCCGTCCTCCT 438

Qy  290 CTTCGTCGGCGCCAAGAACACCATGAACATCATGTCACCAAGTCTGGACGCTGCCGC 349
            |||||                              |||
Db  439 CCCGGTCCTCCGTCCTCCCCGTCCTCCTCCCCGTCCCCGT 498

Qy  350 CGGCGGCGCCTCTTCTACTACCTCTTCGGCTTCGCTTTCGGCTCCCCCTCCCCAACGG 409
            ||||||                           |||
Db  499 CCTCCCCGTCCTCGTCCTCCCCGTCCTCCCCGTCCTCGT 558

Qy  410 CTTCATCGGTAAACATTTCTTCGGCCTCAAGGACATCCTTCATCCTCCTACGACTACAG 469
            ||                               |||
Db  559 CCTCCTCCCGTCCTCCCCGTCCGCGCGGCATCACCAGCGGAAGCAT 618

Qy  470 CTACTTCCTCTACCAATGGGCCTTCGCCATGCGCCGGCATCACCAGCGGAAGCAT 529
            |||||                            |||
Db  619 CGTCCTCCCCGTCCTCGTCCTCCCCGTCCTCCTCCCGTCCTCCC 678

Qy  530 CGCCGAACGCACACAGTTCGTGGCCTATCTCATCATCTCGTCCTCCTTCCTCACCGGCTTCGT 589
            ||||||                           |||
Db  679 CGTCCTCCCCGTCCTCGTCCTCCCCGTCCTCCCCGTCCT 738

Qy  590 CTATCCGGTGGTCTCCCACTGGTTCGGTCCCAGACGGCTGGGCCTCTGCGTTTAAGAT 649
            |||||                            |||
Db  739 CGTCCTCCCCGTCCTCCCCGTCCTCGTCCTCCCCGTCCCCGT 798

Qy  650 CACCGACCGGCTATTTTCCACCGGCGTAATAGACTTCGCCGGTTCGGCGGTAGTCCACAT 709
            |||||                            |||
Db  799 CCTCCCCGTCCTCGTCCTCCCCGTCCTCCCCGTCCTCCCC 858

Qy  710 GGTCGGCGGAATAGCCGCGGCCT 730
            |||||   ||||
Db  859 CGTCCTCGTCCTCCCCGTCCT 879
```

RESULT 4
US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:

RESULT 3
US-09-130-114-2
; Sequence 2, Application US/09130114
; Patent No. 5976807
; GENERAL INFORMATION:
; APPLICANT: Horlick, Robert A.
; APPLICANT: Damaj, Bassam B.
; APPLICANT: Robbins, Alan K.
; TITLE OF INVENTION: Eukaryotic Cells Stably Expressing Genes
; TITLE OF INVENTION: From Multiple Transfected Episomes
; FILE REFERENCE: 0867/1D903US1
; CURRENT APPLICATION NUMBER: US/09/130,114
; CURRENT FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1931
; TYPE: DNA
; ORGANISM: EBNA
US-09-130-114-2

Query Match   5.4%;   Score 91.4; DB 2; Length 1931;
Best Local Similarity 46.7%;   Pred. No. 1.4e-12;
Matches 290; Conservative 0; Mismatches 331; Indels 0; Gaps 0;

```
Qy  319 CCTCCCCGTCCTCCTCCCCGTCCTCCCCGTCCTCGT 378
            ||||||   ||||   |  ||||
Db  230 CCTCTTTTCCGCCTACCTCGTTTTTCTATGCAGCTGGCTTCGCCATGCTCTGCGCCGG 289

Qy  379 CCTCCTCCCGTCCTCGTCCTCCCCGTCCTCCT 438
            ||||||
Db  290 CTTCGTCGGCGCCAAGAACACCATGAACATCATGTCACCAAGTCTGGACGCTGCCGC 349

Qy  439 CCCGGTCCTCCGTCCTCCCCGTCCTCCTCCCCGTCCCCGT 498
            ||||||
Db  350 CGGCGGCGCCTCTTCTACTACCTCTTCGGCTTCGCTTTCGGCTCCCCCTCCCCAACGG 409

Qy  499 CCTCCCCGTCCTCGTCCTCCCCGTCCTCCCCGTCCTCGT 558
            ||||||
Db  410 CTTCATCGGTAAACATTTCTTCGGCCTCAAGGACATCCTTCATCCTCCTACGACTACAG 469

Qy  559 CCTCCTCCCGTCCTCCCCGTCCGCGCGGCATCACCAGCGGAAGCAT 618
            ||||||
Db  470 CTACTTCCTCTACCAATGGGCCTTCGCCATGCGCCGGCATCACCAGCGGAAGCAT 529

Qy  619 CGTCCTCCCCGTCCTCGTCCTCCCCGTCCTCCTCCCGTCCTCCC 678
            ||||||
Db  530 CGCCGAACGCACACAGTTCGTGGCCTATCTCATCATCTCGTCCTCCTTCCTCACCGGCTTCGT 589

Qy  679 CGTCCTCCCCGTCCTCGTCCTCCCCGTCCTCCCCGTCCT 738
            ||||||
Db  590 CTATCCGGTGGTCTCCCACTGGTTCGGTCCCAGACGGCTGGGCCTCTGCGTTTAAGAT 649

Qy  739 CGTCCTCCCCGTCCTCCCCGTCCTCGTCCTCCCCGTCCCCGT 798
            ||||||
Db  650 CACCGACCGGCTATTTTCCACCGGCGTAATAGACTTCGCCGGTTCGGCGGTAGTCCACAT 709

Qy  799 CCTCCCCGTCCTCGTCCTCCCCGTCCTCCCCGTCCTCCCC 858
            ||||||
Db  710 GGTCGGCGGAATAGCCGCGGCCT 730

Qy  859 CGTCCTCGTCCTCCCCGTCCT 879
            |||||   ||||
Db      CCAACTTCTCGCCCAAACACCACAGACGCCTCGCCGCCTCCCTTATCTGCGGCCA 169
```

```
Qy  110 CCAACTTCTCGCCCAAACACCACAGACGCCTCGCCGCCTCCCTTATCTGCGGCCA 169
            ||||| |||  ||||   ||  ||||
Db  259 CCAACCTTGTCCTCGTCCTCCCCGTCCTCCCCGTCCTCCCCGTCCTCGTCCT 318

Qy  170 TTTCGCCGCCTGGACAGCAGCAAGTTCGTCGACACGGCCTTCGCCGTCGACAAACCTACCT 229
            |||||
Db  319 CCTCCCCGTCCTCGTCCTCCCCGTCCTCCCCGTCCTCCCCGTCCTCCTCCCCGTCCTCCT 378
```

; OTHER INFORMATION: having a 120 repeat of ACG flanked by fixed
; OTHER INFORMATION: fragments having NcoI restriction sites.
US-09-197-649-7

Query Match   4.5%;   Score 76.6;   DB 3;   Length 390;
Best Local Similarity   50.1%;   Pred. No. 2.7e-09;
Matches 190;   Conservative   0;   Mismatches 189;   Indels   0;   Gaps   0;

```
Qy    163 GCGGGCCATTTGCCGCCGTGACAGCAAGTTCGTCGACACGGCCTTCGCGTCGACAACA 222
          ||||| ||||||||||||| |||||| || |||||||||||||||||||||
Db    379 GCAACCATGGCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 320

Qy    223 CCTACCTCCTCCTTTCCGCCTACCTCGTTTTCTATGCAGCTCGTCGTCGTCGCCATGCTCT 282
          ||| |
Db    319 TCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGGTCG 260

Qy    283 GCGCCGGCTCCGTCCGCGCCAAGAACACCATGAACATGCTCACCAACGTCTCCACACG 342
          ||| ||||||||||||||||||
Db    259 TCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 200

Qy    343 CTGCGCCGCGGCGGCCTTCTTCTACTACCTCTTCGGCTTCGCCTTCTCGCTTTCGGCTCCCCT 402
          || |||||||||||||||
Db    199 TCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 140

Qy    403 CCAACGGCTTCATCGGTAAACATTTCTTCGGCTTCAAGGACAATCCCTTCGTCGTCGTCCCTACG 462
          ||||||||||| ||
Db    139 TCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 80

Qy    463 ACTACAGCTACTTCCTCTACCAATGGCCTTCGCCATCGCCGCGCCGGCATCACCAGCG 522
          ||||||||||||
Db     79 TCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 20

Qy    523 GAAGCATCGCGAACGCAC 541
          |||||||
Db     19 TCGTCGTCGTCCATGGCCC 1
```

RESULT 6
US-09-249-585A-2/c
; Sequence 2, Application US/09249585A
; Patent No. 6417002
; GENERAL INFORMATION:
; APPLICANT: Horlick, Robert
; TITLE OF INVENTION: METHOD FOR MAINTENANCE AND SELECTION OF EPISOMES
; FILE REFERENCE: 0867/0D905
; CURRENT APPLICATION NUMBER: US/09/249,585A
; CURRENT FILING DATE: 1999-02-11
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 1926
; TYPE: DNA
; ORGANISM: Epstein Barr Virus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1926)
; OTHER INFORMATION: coding strand of EBNA-1 DNA
US-09-249-585A-2

Query Match   4.5%;   Score 76;   DB 4;   Length 1926;
Best Local Similarity   47.2%;   Pred. No. 7.8e-09;
Matches 232;   Conservative   0;   Mismatches 260;   Indels   0;   Gaps   0;

```
Qy    137 CGCTCCGCGCGCGCCTTCCCTTATCTGCGGCCATTTCGCCGCCGTGGACAGCAAGTTCGT 196
          || ||||||||||||||||
Db    745 CTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCTCCTGCTCCTGCCC 686

Qy    197 CGACACGGCCTTCGCCGTCGACAACACCACCTCCTCTTTTCCGGCTCACCTCGTTTTTC 256
          || |||||||||||||||||||
Db    685 CTCCTGCCCCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCCC 626

Qy    257 TATGCAGCTCGGCTTCGCCATGCTCTGCGGCGGCTTCGCCGCCGCGGCCAAGAACACCATGAA 316
          || |||
```

INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-F18
US-08-232-463-14

Query Match   4.8%;   Score 82;   DB 1;   Length 7218;
Best Local Similarity   6.7%;   Pred. No. 5e-10;
Matches   28;   Conservative 241;   Mismatches 151;   Indels   0;   Gaps   0;

```
Qy    175 CCGCCGTGACAGCAAGTTCGTCGACACGGCCTTCGCGTCGACAACCTACCTCCTCT 234
          ::|||| :::::: :: |::: :::::::::
Db   1033 CCGAGCTTGGCTGCAGGTGCAGGGAGCTTGCGATYYYYYYYYYYYYYYYYYYYYYY 1092

Qy    235 TTTCCGCCTACCTCGTTTTTCTATGCAGCTCGGCTTCGCCATGCTCTGCGCCGGCGGCTCCG 294
          :: ::: :::: :::::::::::YYYYYYYYYYYYYYYYYYYYYYYYYYYYYY
Db   1093 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1152

Qy    295 TCGGCGCCAAGAACACCATGAACATGCTCACCAACGTCTCCGGACGGCTGCCGCCGGCG 354
          ::: ::::: :: :::::::::: :: :: ::: :
Db   1153 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1212

Qy    355 GCCTCTTCTACTACCTCTTCGGCTTCGCCTTCGCTTTCGGCTTCCCCCTCCCCCCACGGCTTCA 414
          ::: :::::: ::::::::: :::: :::::::::::YYYYYYYYYYY ::::
Db   1213 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1272

Qy    415 TCGGTAAACATTTCTTCGGCTTCAAGGACAATCCCTTCATCCTCTACGACTACAGCTACT 474
          ::: ::: : :::::::::: :::::: ::::::::::::YYYYYYYYYYYY
Db   1273 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1332

Qy    475 TCCTCTACCAATGGGCCTTCGCCATCGCCGCGCCGGCATCACCAGCGGAAGCATCGCGCG 534
          ::::::: :: ::::: :: :: ::: :: ::: :: ::: :
Db   1333 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1392

Qy    535 AACGGCACACAGTTCGTGGCCTATCTCATCTACTACCTCCTCCTTCCTCCTCACCGGCTTCGTCATC 594
          ::: ::: :::::::::: ::::::::::::::::
Db   1393 YYYYYYYYYYYYYYYYYYYYYYYYYYYYGTACCCAAATTCTTCTATC 1452
```

RESULT 5
US-09-197-649-7/c
; Sequence 7, Application US/09197649
; Patent No. 6194550
; GENERAL INFORMATION:
; APPLICANT: Gold, Larry
; APPLICANT: Tuerk, Craig
; APPLICANT: Pribnow, David
; APPLICANT: Smith, Jonathan D.
; TITLE OF INVENTION: Systematic Polypeptide Evolution by Reverse Translation
; FILE REFERENCE: NEX02/C1-CON
; CURRENT APPLICATION NUMBER: US/09/197,649
; CURRENT FILING DATE: 1998-11-23
; EARLIER APPLICATION NUMBER: 07/829,461
; EARLIER APPLICATION FILING DATE: 1992-01-31
; EARLIER APPLICATION NUMBER: 07/739,055
; EARLIER APPLICATION FILING DATE: 1991-08-01
; EARLIER APPLICATION NUMBER: 07/561,968
; EARLIER APPLICATION FILING DATE: 1990-08-02
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 390
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Sequence

Left column

```
Db  625 CTCCTGCCCCTCCTGCTCCTGCTCCTGCGCCCTCCTGCCCTCTCCTGCGCCCTCCTGCCC 566
Qy  317 CATCATGCTCACCAACGTCCTGGAGCGCTGCGGCGGCCTTCTTCTACTACCTCTTCGG 376
Db  565 CTCCTCCTGCTCCTGCCCCTCCTGCTCCTGCGCCCTCCTGCTCCTGCTCCTGCTCCTG 506
Qy  377 CTTCGGCCTTCGCTTTCGGCTCCCCTCCAACGGGCTTCATCGGTAAACATTTCTTCGGCCT 436
Db  505 CCCCTTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCTCCTGCTCCTGCTCCTGCTC 446
Qy  437 CAAGGACATCCCTTCATCCTCTACGACTACAGCTACTTCCTCTACCAATGGGCCTTCGC 496
Db  445 CTGCCCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCTCCTGCTCCTGCTCCTG 386
Qy  497 CATCGCGCCGCCGGCATCACCAGCGGAAGCATCGCCGAACGCACACAGTTCGTGGCCTA 556
Db  385 CTCCTCCTGCTCCTGCCCCTCCTGCGCCCTCCTGCGCTATCCGGTGCTCCCACTGTTCTG 326
Qy  557 TCTCATCTACTCCTTCCTTCCTCACCGGCTTCGTCTATCGGTGTCCCGTGGTCTCCACTGGTTCTG 616
Db  325 CCCCTCCTGCTCCTGCTCCTGCCCCTCCTGCTCCGCTCCGCTCCTGCTCCTGCTCCTG 266
Qy  617 GTCCCCAGACGG 628
Db  265 TTCCACCGTGGG 254
```

RESULT 7
US-09-410-399-3/c
; Sequence 3, Application US/09410399
; Patent No. 6482587
; GENERAL INFORMATION:
;   APPLICANT: Robertson, Erle S.
;   APPLICANT: Cotter, Murray A.
;   TITLE OF INVENTION: Methods to Inhibit or Enhance the Binding of Viral DNA
;   TITLE OF INVENTION: to Genomic Host DNA
;   FILE REFERENCE: UM-03778
;   CURRENT APPLICATION NUMBER: US/09/410,399
;   CURRENT FILING DATE: 1999-10-01
;   NUMBER OF SEQ ID NOS: 6
;   SOFTWARE: PatentIn Ver. 2.0
;   SEQ ID NO 3
;   LENGTH: 1926
;   TYPE: DNA
;   ORGANISM: Epstein-Barr virus
US-09-410-399-3

Query Match   4.5%; Score 76; DB 4; Length 1926;
Best Local Similarity 47.2%; Pred. No. 7.8e-09;
Matches 232; Conservative 0; Mismatches 260; Indels 0; Gaps 0;

```
Qy  137 CGCCTCCGCGCGCCGCCTCCCTTATCTGCGGCGCCATTTCGCGCCGCCGGTGGACAGCAAGTTCGT 196
Db  745 CTCCTGCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCTCCTGCTCCTGCTCCTGCCC 686
Qy  197 CGACACGGCCTTCGGCGTCGACAACACTACCTCCTCCTTTTCCGCCTACCTCGTTTTTTC 256
Db  685 CTCCTGCCCCTCCTGCCTCCTGCTCCTGCTCCTGCTCCTGCCCCTCCTGCCC 626
Qy  257 TATGCAGCTCGGCTTCGGCCATGCTTCTGCGCGGCCTCTGCGCCGGCGCAGAACACCATGAA 316
Db  625 CTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCTCCTGCTCCTGCTCCTGCCC 566
Qy  317 CATCATGCTCACCAACGTCCTGGAGCGTCCTGGAGCGCCTTCTTCTACTACCTCTTCGG 376
Db  565 CTCCTCCTGCTCCTGCCCCTCCTGCTCCTGCGCCCTCCTGCTCCTGCTCCTGCTCCTG 506
Qy  377 CTTCGGCCTTCGCTTTCGGCTCCCCTCCAACGGGCTTCATCGGTAAACATTTCTTCGGCCT 436
Db  505 CCCCTTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCTCCTGCTCCTGCTCCTGCTC 446
Qy  437 CAAGGACATCCCTTCATCCTCTACGACTACAGCTACTTCCTCTACCAATGGGCCTTCGC 496
```

Right column

```
Db  445 CTGCCCCTCCTGCTCCTGCTCCTGCGCCCTCCTGCCCCTCCTGCTCCTGCTCCTGCCC 386
Qy  497 CATCGCGCCGCCGGCATCACCAGCGGAAGCATCGCCGAACGCACACAGTTCGTGGCCTA 556
Db  385 CTCCTCCTGCTCCTGCCCCTCCTGCGCCCTCCTGCGCCCTCCTGCTCCTGCTCCTG 326
Qy  557 TCTCATCTACTCCTTCCTTCCTCACCGGCTTCGTCTATCGGTGTCCCCACTGGTTCTG 616
Db  325 CCCCTTCCTGCTCCTGCTCCTGCCCCTCCCGCCCCTGCTCCGCTCCTGCTCCTGCTCCTG 266
Qy  617 GTCCCCAGACGG 628
Db  265 TTCCACCGTGGG 254
```

RESULT 8
US-09-050-863-2/c
; Sequence 2, Application US/09050863
; Patent No. 6114111
; GENERAL INFORMATION:
;   APPLICANT: Lao, Ying
;   APPLICANT: Hiang, Betty
;   APPLICANT: Payan, Don
;   TITLE OF INVENTION: Mammalian Protein Interaction Cloning
;   TITLE OF INVENTION: System
;   NUMBER OF SEQUENCES: 5
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
;   STREET: 4 Embarcadero Center, Suite 3400
;   CITY: San Francisco
;   STATE: CA
;   COUNTRY: USA
;   ZIP: 94111-4187
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/050,863
;   FILING DATE: 30-MAR-1998
;   CLASSIFICATION:
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Silva, Robin M.
;   REGISTRATION NUMBER: 38,304
;   REFERENCE/DOCKET NUMBER: A-65638/DJB/RMS
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (415) 781-1989
;   TELEFAX: (415) 949-8711
;   INFORMATION FOR SEQ ID NO: 2:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 2580 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: unknown
;   TOPOLOGY: unknown
;   MOLECULE TYPE: DNA
US-09-050-863-2

Query Match   4.5%; Score 76; DB 3; Length 2580;
Best Local Similarity 47.2%; Pred. No. 8.9e-09;
Matches 232; Conservative 0; Mismatches 260; Indels 0; Gaps 0;

```
Qy  137 CGCCTCCGCGCGCCGCCTCCCTTATCTGCGGCGCCATTTCGCGCCGCCGGTGGACAGCAAGTTCGT 196
Db 1128 CTCCTGCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCTCCTGCTCCTGCTCCTGCTCCGCCC 1069
Qy  197 CGACACGGCCTTCGGCGTCGACAACACTACCTCCTCCTTTTCCGCCTACCTCGTTTTTTC 256
Db 1068 CTCCTGCCCCTCCTGCTCCTGCTCCTGCTCCTGCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCCC 1009
Qy  257 TATGCAGCTCGGCTTCGGCCATGCTTCGGCGCGGCCTCTGCGCCGGCCAAGAACACCATGAA 316
Db 1008 CTCCTGCCCCTCCTGCTCCTGCTCCTGCCCCTCCTGCTCCTGCTCCTGCTCCTGCGCCCCTCCTGCCC 949
```

Query Match   4.5%;   Score 76;   DB 4;   Length 2580;
Best Local Similarity   47.2%;   Pred. No. 8.9e-09;
Matches 232;   Conservative   0;   Mismatches 260;   Indels   0;   Gaps   0;

```
Qy   137  CGCTCCGCCGCGGCCTCCTTATCTGCGGGCCATTCGCGCCGTGGACAGCAAGTTCGT  196
               |||| |||||||| |||||||||||||||||||||
Db  1128  CTCCTGCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCTCCTCCTGCCCC  1069

Qy   197  CGACACGGCCTTCGCCGTCGACAACACTACCTCCTCTTTCCGCTACCTCGTTTTTTC  256
               ||||| |||||||||||||||||||||
Db  1068  CTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCTCCTGCCCC  1009

Qy   257  TATGCAGCTCGGCTTCGCCATGCTCGTGCCGCCGCTCGGTCCGGCCAAGAACACCATGAA  316
               ||| ||||||||||||||
Db  1008  CTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCCCC  949

Qy   317  CATCATGCTCACCAACGTCCTGGACGCTGCGCGCCGGCGGCCTCTTCTACTACCTCTTCGG  376
               ||| ||||||||||||||
Db   948  CTCCTGCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCTCCTGCCCCTCCTG  889

Qy   377  CTTCGGCTTCGTTTTCGGCTCCCCCCTCCAAGGCTTCATCGGTAAACATTTCTTCGGCCT  436
               ||||||||||||
Db   888  CCCCTCCTGCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCTCCTGCTCCTGCTC  829

Qy   437  CAAGGACATCCCTTCATCCTCTACGACTACAGCTACTTCCTCTACCAATGGCGCTTCGC  496
               |||||||||||||
Db   828  CTGCCCCTCCTGCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCTCCTGCCCC  769

Qy   497  CATCGCCGGCCGCCGGCATCACCAGCGGAAGCATCGCCGAACGCACACAGTTCGTGGGCCTA  556
               |||||||||||||||
Db   768  CTCCTGCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCTCCTGCTCCTGCTCCTG  709

Qy   557  TCTCATCTACTCCTCCTTCCTCACCGGCTTCGTCATCCGGTTGGTCTCCACTGGTTCTG  616
               ||||||||||||||
Db   708  CCCCTCCTGCTCCTGCTCCTGCCCCTCCTGCCCCTCCGCCTCCGCCTCCTGCTCCTCCTG  649

Qy   617  GTCCCCAGACGG  628
               ||| |||
Db   648  TTCCACCGTGGG  637
```

RESULT 10
US-09-130-114-1
; Sequence 1, Application US/09130114
; Patent No. 5976807
; GENERAL INFORMATION:
;   APPLICANT: Horlick, Robert A.
;   APPLICANT: Damaj, Bassam B.
;   APPLICANT: Robbins, Alan K.
;   TITLE OF INVENTION: Eukaryotic Cells Stably Expressing Genes
;   TITLE OF INVENTION: From Multiple Transfected Episomes
;   FILE REFERENCE: 0867/1D903US1
;   CURRENT APPLICATION NUMBER: US/09/130,114
;   CURRENT FILING DATE: 1998-08-06
;   NUMBER OF SEQ ID NOS: 36
;   SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
;   LENGTH: 5452
;   TYPE: DNA
;   ORGANISM: VEBNA
US-09-130-114-1

Query Match   4.5%;   Score 76;   DB 2;   Length 5452;
Best Local Similarity   47.2%;   Pred. No. 1.2e-08;
Matches 232;   Conservative   0;   Mismatches 260;   Indels   0;   Gaps   0;

```
Qy   137  CGCTCCGCCGCGGCCTCCTTATCTGCGGGCCATTCGCGCCGTGGAGCAGCAAGTTCGT  196
               |||| |||||||| |||||||||||||||||||||
Db  1677  CTCCTGCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCTCCTCCTGCCCC  1736

Qy   197  CGACACGGCCTTCGCCGTCGACAACACTACCTCCTCTTTCCGCTACCTCGTTTTTTC  256
               ||||| |||||||||||||||||||||
Db  1737  CTCCTGCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCTCCTGCCCC  1796
```

```
ly   317  CATCATGCTCACCAACGTCCTGGACGCTGCGCGGCCGGCGGCCTCTTCTACTACCTCTTCGG  376
               ||| ||||||||||||||
lb   948  CTCCTGCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCTCCTGCCCCTCCTG  889

ly   377  CTTCGGCTTCGTTTTCGGCTCCCCCCTCCAACGGCTTCATCGGTAAACATTTCTTCGGCCT  436
               |||| |||||||||||
lb   888  CCCCTCCTGCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCTCCTGCTC  829

ly   437  CAAGGACATCCCTTCATCCTCTACGACTACAGCTACTTCCTCTACCAATGGCGCTTCGC  496
               |||||||||||||
lb   828  CTGCCCCTCCTGCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCTCCTGCCCC  769

ly   497  CATCGCCGGCCGCCGGCATCACCAGCGGAAGCATCGCCGAACGCACACAGTTCGTGGGCCTA  556
               |||||||||||||||
lb   768  CTCCTGCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCTCCTGCTCCTGCTCCTG  709

ly   557  TCTCATCTACTCCTCCTTCCTCACCGGCTTCGTCATCCGGTTGGTCTCCACTGGTTCTG  616
               ||||||||||||||
lb   708  CCCCTCCTGCTCCTGCTCCTGCCCCTCCTGCCCCTCCGCCTCCGCCTCCTGCTCCTCCTG  649

ly   617  GTCCCCAGACGG  628
               ||| |||
b    648  TTCCACCGTGGG  637
```

ESULT 9
S-09-359-081-2/c
; Sequence 2, Application US/09359081
; Patent No. 6316223
; GENERAL INFORMATION:
;   APPLICANT: Lao, Ying
;   Hiang, Betty
;   Payan, Don
;   TITLE OF INVENTION: Mammalian Protein Interaction Cloning
;   System
;   NUMBER OF SEQUENCES: 5
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
;   STREET: 4 Embarcadero Center, Suite 3400
;   CITY: San Francisco
;   STATE: CA
;   COUNTRY: USA
;   ZIP: 94111-4187
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/359,081
;   FILING DATE: 22-Jul-1999
;   CLASSIFICATION: <Unknown>
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 09/050,863
;   FILING DATE: <Unknown>
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Silva, Robin M.
;   REGISTRATION NUMBER: 38,304
;   REFERENCE/DOCKET NUMBER: A-65638/DJB/RMS
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (415) 781-1989
;   TELEFAX: (415) 949-8711
; INFORMATION FOR SEQ ID NO: 2:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 2580 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: unknown
;   TOPOLOGY: unknown
;   MOLECULE TYPE: DNA
;   SEQUENCE DESCRIPTION: SEQ ID NO: 2:
S-09-359-081-2

```
Qy    257  TATGCAGCTCGGCTTCGCCATGCTCGCGCCGGCTCCGCGCCGCCAAGAACACCATGAA  316
                |||||||||||||||||||
Db   1797  CTCCTGCCCCCTCCTGCTCTCTGCCCCCTCCTGCTCTCTGCCCCCTCCTGCTCTCTG  1856

Qy    317  CATCATGCTCACCAACGTCCTGGACGTCGCGCCGGGCCCTTCTTCTACTACCTCTTCGG  376
                |||||||||||||||||||
Db   1857  CTCCTGCTCCTGCTGCCCCTCCTGCTCTCTGCCCCCTCCTGCTCTCTGCTCCTCCTG  1916

Qy    377  CTTGCCTTCGCTTTCGGCTCCCCCTCCAAGGCTTCATCGGTAAACATTTCTTCGGCCT  436
                |||||||||||||||||||
Db   1917  CCCCTCCTGCCCCCTCCTGCTCTCTGCCCCCTCCTGCTCTCTGCCCCCTCCTGCTC  1976

Qy    437  CAAGGACATCCCTCATCCTCCTACGACTACAGCTACTTCCTCTACCAATGGGCCTTCGC  496
                |||||||||||||||||||
Db   1977  CTGCCCCCTCCTGCTCTCTGCCCCCTCCTGCTCTCTGCCCCCTCCTGCTCTCTGCCCC  2036

Qy    497  CATCGCGCCGCGGCATCACCAGCGGAAGCATCGCCGAACGCACACAGTTCGTGGCCTA  556
                |||||||||||||||||||
Db   2037  CTCCTGCTCCTGCTGCCCCTCCTGCTCTCTGCCCCCTCCTGCTCTCTGCTCCTCCTG  2096

Qy    557  TCTCATCTACTCCTTCTTCGTTCATCCGGTGGTGTCCCACTGGTTCTG  616
                |||||||||||||||||||
Db   2097  CCCCTCCTGCTCTCTGCCCCCTCCTGCTCCCCGTCTCCTGCTCTCTG  2156

Qy    617  GTCCCAGACGG  628
                |||||||
Db   2157  TTCCACCGTGGG  2168
```

RESULT 11
US-09-647-344A-14
; Sequence 14, Application US/09647344A
; Patent No. 6586180
; GENERAL INFORMATION:
; APPLICANT: Ruffner, Duane E.
; APPLICANT: Pierce, Michael L.
; APPLICANT: Chen, Zhidong
; TITLE OF INVENTION: Directed Antisense Libraries
; FILE REFERENCE: T6678.PCT.US
; CURRENT APPLICATION NUMBER: US/09/647,344A
; CURRENT FILING DATE: 2000-12-04
; PRIOR APPLICATION NUMBER: PCT/US99/06742
; PRIOR FILING DATE: 1999-03-28
; NUMBER OF SEQ ID NOS: 50
; SEQ ID NO 14
; LENGTH: 8705
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: pShuttle
US-09-647-344A-14

```
Query Match      4.5%;  Score 76;  DB 4;  Length 8705;
Best Local Similarity   47.2%;  Pred. No. 1.5e-08;
Matches 232;  Conservative   0;  Mismatches 260;  Indels   0;  Gaps   0;

Qy    137  CGCCTCCGCGCGCCGCTCCCTTATCTGGGCGCCATTCGCCGCGTGGACAGCAAGTTCGT  196
                ||||||||||
Db   7542  CTCCTGCTCCTGCTCTCTGCCCCCTCCTGCTCTCTGCCCCCTCCTGCTCTCTGCTCCGCC  7601

Qy    197  CGACACGGCCTTCGCCGTCGACAACACCTACCTCCTCTTTTCCGCCTACCTCGTTTTTC  256
                |||||||||||||
Db   7602  CTCCTGCCCCCTCCTGCTCTCTGCCCCCTCCTGCTCTCTGCCCCCTCCTGCTCCTGCCC  7661

Qy    257  TATGCAGCTCGGCTTCGCCATGCTCGCGCCGGCTCCGCGCCGCCAAGAACACCATGAA  316
                |||||||||||||||||||
Db   7662  CTCCTGCCCCCTCCTGCTCTCTGCCCCCTCCTGCTCTCTGCCCCCTCCTGCTCTGCCC  7721

Qy    317  CATCATGCTCACCAACGTCCTGGACGTCGCGCCGGGCCCTTCTTCTACTACCTCTTCGG  376
                |||||||||||||||||||
Db   7722  CTCCTGCTCCTGCTGCCCCTCCTGCTCTCTGCCCCCTCCTGCTCTCTGCTCCTCCTG  7781
```

```
Qy    377  CTTCGCCTTCGCTTTCGGCTCCCCCTCCAAGGCTTCATCGGTAAACATTTCTTCGGCCT  436
                |||||||||||||||||||
Db   7782  CCCCTCCTGCCCCCTCCTGCTCTCTGCCCCCTCCTGCTCTCTGCCCCCTCCTGCTC  7841

Qy    437  CAAGGACATCCCTCATCCTCCTACGACTACAGCTACTTCCTCTACCAATGGGCCTTCGC  496
                |||||||||||||||||||
Db   7842  CTGCCCCCTCCTGCTCTCTGCCCCCTCCTGCTCTCTGCCCCCTCCTGCTCTCTGCCCC  7901

Qy    497  CATCGCGCCGCGGCATCACCAGCGGAAGCATCGCCGAACGCACACAGTTCGTGGCCTA  556
                |||||||||||||||||||
Db   7902  CTCCTGCTCCTGCTGCCCCTCCTGCTCTCTGCCCCCTCCTGCTCTCTGCTCCTCCTG  7961

Qy    557  TCTCATCTACTCCTTCTTCGTTCATCCGGTGGTGTCCCACTGGTTCTG  616
                |||||||||||||||||||
Db   7962  CCCCTCCTGCTCTCTGCCCCCTCCTGCTCCCCGTCTCCTGCTCTCTG  8021

Qy    617  GTCCCAGACGG  628
                |||||||
Db   8022  TTCCACCGTGGG  8033
```

RESULT 12
US-08-910-647-1/c
; Sequence 1, Application US/08910647
; Patent No. 6251433
; GENERAL INFORMATION:
; APPLICANT: Zuckermann et al.
; TITLE OF INVENTION: Compositions and Methods for
; TITLE OF INVENTION: Polynucleotide Delivery
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/910,647
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Fujita, Sharon M.
; REGISTRATION NUMBER: 38,459
; REFERENCE/DOCKET NUMBER: 1218.002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 923-2706
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9600 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-910-647-1

```
Query Match      4.5%;  Score 76;  DB 3;  Length 9600;
Best Local Similarity   47.2%;  Pred. No. 1.6e-08;
Matches 232;  Conservative   0;  Mismatches 260;  Indels   0;  Gaps   0;

Qy    137  CGCCTCCGCGCGCCGCTCCCTTATCTGGGCGCCATTCGCCGCGTGGACAGCAAGTTCGT  196
                ||||||||||
Db   1174  CTCCTGCTCCTGCTCTCTGCCCCCTCCTGCTCTCTGCCCCCTCCTGCTCTCTGCTCCGCC  1115

Qy    197  CGACACGGCCTTCGCCGTCGACAACACCTACCTCCTCTTTTCCGCCTACCTCGTTTTTC  256
                |||||||||||||
Db   1114  CTCCTGCCCCCTCCTGCTCTCTGCCCCCTCCTGCTCTCTGCCCCCTCCTGCTCCTGCCCC  1055
```

US-09-620-925-1

Query Match   4.5%;   Score 76; DB 4; Length 9600;
Best Local Similarity 47.2%;   Pred. No. 1.6e-08;
Matches   232; Conservative   0; Mismatches 260; Indels   0; Gaps   0;

```
Qy    137 CGCTCCGCCGCGCCTTCCCTTATCTGCGGCCATTTCGCCGCCGTGGACAGCAAGTTCGT 196
              ||||| ||    || ||   |||
Db   1174 CTCCTGCTCCTGCCCCTCCTGTCTCCTGCGCCCCTCCTGCTCCTGCCCCTCCTGCCCC 1115

Qy    197 CGACACGGCCTTCGCCGTCGACAACACTACCTCCTCTTTTCCGCCTACCTCGTTTTTC 256
              ||||| ||  || ||
Db   1114 CTCCTGCCCCTCCTGCTCCTGCCCCTCCTGTCCTGCCCCTCCTGCCCCTCCTGCCC 1055

Qy    257 TATGCAGCTCGGCTTCGCCATGCTCTGCGCGCCTCCGGCGCCAAGAACACCATGAA 316
              || ||  ||
Db   1054 CTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCCC 995

Qy    317 CATCATGCTCACCAACGTCCTGGACGGTCGCGCCGGCGCCTCTTCTACTACCTCTTCGG 376
              ||| ||      ||
Db    994 CTCCTGCTCCTGCCCCTCCTGTCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCCCTCCTG 935

Qy    377 CTTCGGCTTTCGGCTTCCCCTCAACGGCTTCATCGGTAAACATTTCTTCGGCCT 436
              ||| ||    |||
Db    934 CCCCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCCCCTCCTGCTC 875

Qy    437 CAAGGACATCCCTTCATCCTCCTAGACTACAGCTACTTCCTCTACCAATGGGCCTTCGC 496
              ||  ||       |||
Db    874 CTGCCCCTCCTGCCCCTCCTGCCCCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCCCC 815

Qy    497 CATCGCCGCCGGCGATCACCAGCGGAAGCATCGCCGAACGCACACAGTTCGTCGGCCTA 556
              ||| ||    ||
Db    814 CTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCCCCTCCTGCTCCTGCCCCTCCTCCTG 755

Qy    557 TCTCATCTACTCCTCCTTCCTCACCGGCTTCGTCTATCCGGTGGTCTCCCACTGGTTCTG 616
              ||| ||    |||
Db    754 CCCCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCCCCGGCTCCTGCGTCCTGCTCCTG 695

Qy    617 GTCCCCAGACGG 628
              ||||
Db    694 TTCCACCGTGGG 683
```

RESULT 13
US-09-620-925-1/c
Sequence 1, Application US/09620925
Patent No. 6468996
GENERAL INFORMATION:
    APPLICANT: Zuckermann et al.
    TITLE OF INVENTION: Compositions and Methods for
                       Polynucleotide Delivery

    NUMBER OF SEQUENCES: 4
    CORRESPONDENCE ADDRESS:
        ADDRESSEE: Chiron Corporation
        STREET: 4560 Horton Street
        CITY: Emeryville
        STATE: California
        COUNTRY: U.S.A.
        ZIP: 94608-2916
    COMPUTER READABLE FORM:
        MEDIUM TYPE: Floppy disk
        COMPUTER: IBM PC compatible
        OPERATING SYSTEM: PC-DOS/MS-DOS
        SOFTWARE: PatentIn Release #1.0, Version #1.30
    CURRENT APPLICATION DATA:
        APPLICATION NUMBER: US/09/620,925
        FILING DATE: 21-Jul-2000
        CLASSIFICATION: <Unknown>
    PRIOR APPLICATION DATA:
        APPLICATION NUMBER: 08/910,647
        FILING DATE: <Unknown>
    ATTORNEY/AGENT INFORMATION:
        NAME: Fujita, Sharon M.
        REGISTRATION NUMBER: 38,459
        REFERENCE/DOCKET NUMBER: 1218.002
    TELECOMMUNICATION INFORMATION:
        TELEPHONE: (510) 923-2706
        TELEFAX: (510) 655-3542
    INFORMATION FOR SEQ ID NO: 1:
        SEQUENCE CHARACTERISTICS:
            LENGTH: 9600 base pairs
            TYPE: nucleic acid
            STRANDEDNESS: single
            TOPOLOGY: linear
        MOLECULE TYPE: DNA (genomic)
        SEQUENCE DESCRIPTION: SEQ ID NO: 1:

RESULT 14
US-07-884-811-15/c
Sequence 15, Application US/07884811
Patent No. 5316921
GENERAL INFORMATION:
    APPLICANT: Godowski, Paul J. Lokker, Nathalie A. Mark, Melanie R.
    TITLE OF INVENTION: SINGLE-CHAIN HEPATOCYTE GROWTH FACTOR VARIANTS
    NUMBER OF SEQUENCES: 21
    CORRESPONDENCE ADDRESS:
        ADDRESSEE: Genentech, Inc.
        STREET: 460 Point San Bruno Blvd
        CITY: South San Francisco
        STATE: California
        COUNTRY: USA
        ZIP: 94080
    COMPUTER READABLE FORM:
        MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
        COMPUTER: IBM PC compatible
        OPERATING SYSTEM: PC-DOS/MS-DOS
        SOFTWARE: patin (Genentech)
    CURRENT APPLICATION DATA:
        APPLICATION NUMBER: US/07/884,811
        FILING DATE: 19920518
        CLASSIFICATION: 530
    PRIOR APPLICATION DATA:
        APPLICATION NUMBER:
        FILING DATE:
    ATTORNEY/AGENT INFORMATION:
        NAME: Dreger, Ginger R.
        REGISTRATION NUMBER: 33,055

```
          SOFTWARE: patin (Genentech)
  CURRENT APPLICATION DATA:
     APPLICATION NUMBER: US/07/885,971
     FILING DATE: 1992O518
     CLASSIFICATION: 530
  PRIOR APPLICATION DATA:
     APPLICATION NUMBER:
     FILING DATE:
  ATTORNEY/AGENT INFORMATION:
     NAME: Dreger, Ginger R.
     REGISTRATION NUMBER: 33,055
     REFERENCE/DOCKET NUMBER: 779
  TELECOMMUNICATION INFORMATION:
     TELEPHONE: 415/225-3216
     TELEFAX: 415/952-9981
     TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 15:
  SEQUENCE CHARACTERISTICS:
     LENGTH: 10596 bases
     TYPE: NUCLEIC ACID
     STRANDEDNESS: single
     TOPOLOGY: linear
US-07-885-971-15
```

```
Query Match          4.5%;   Score 76;   DB 1;   Length 10596;
Best Local Similarity  47.2%;   Pred. No. 1.7e-08;
Matches 232;   Conservative  0;   Mismatches 260;   Indels  0;   Gaps  0;

Qy  137 CGCCTCCGCCGCGGCGCCTCCTTATCTGGGCCATTTCGCGCCGTGGACAGCAAGTTCGT 196
             ||||||
Db 2708 CTCCTGCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCCC 2649

Qy  197 CGACACGGCCTTCGCGTGACAACACCTACCTCCTCTTTCCGCTACCTCGTTTTTC 256
             ||||
Db 2648 CTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCCC 2589

Qy  257 TATGCAGCTCGGCTTCGCCATGCTCTGCGCCGGCTCCTGCGCCGCCAAGAACACCATGAA 316
             ||||
Db 2588 CTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCGCCCCTCCTGCCC 2529

Qy  317 CATCATGCTCACCAAGTCCTGAGACGTCCGGCGCGGCCTCTTCTACTACCTCCTCGG 376
             ||||||
Db 2528 CTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCTCCTG 2469

Qy  377 CTTCGGCTTGGCTTTCGGCTCCCCCCTCCAAGGGCTTCATCGGTAAACATTTCTTCGGGCCT 436
             |||||
Db 2468 CCCCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCTCCTGCTC 2409

Qy  437 CAAGGACATCCCTTCATCCTCTACGACTACAGCTACTTCCTCTACCAATGGGCCTTCGC 496
             |||||
Db 2408 CTGCCCCTCCTGCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCTCCTGCCC 2349

Qy  497 CATCGCCGCGCCGCGGCATCACCAGCGGAAGCATCGCCGAAGCACACAGTTCGTGGCCTA 556
             ||||||
Db 2348 CTCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCTCCTGCTGGTCCTCCTGCTCCTG 2289

Qy  557 TCTCATCTACTCCTTCCTTCCTCACCGGCTTCGTCTATCCGGTGGTCTTCCCACTGGTTCTG 616
             |||||
Db 2288 CCCCTCCTGCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCTCCTGCTCCTCCTG 2229

Qy  617 GTCCCCAGACGG 628
             |||||
Db 2228 TTCCACCGTGGG 2217
```

```
RESULT 15
US-07-885-971-15/c
Sequence 15, Application US/07885971
Patent No. 5328837
GENERAL INFORMATION:
  APPLICANT: Godowski, Paul J. Lokker, Nathalie A. Mark, Melanie R.
  TITLE OF INVENTION: HEPATOCYTE GROWTH FACTOR PROTEASE DOMAIN VARIANTS
  NUMBER OF SEQUENCES: 21
  CORRESPONDENCE ADDRESS:
     ADDRESSEE: Genentech, Inc.
     STREET: 460 Point San Bruno Blvd
     CITY: South San Francisco
     STATE: California
     COUNTRY: USA
     ZIP: 94080
  COMPUTER READABLE FORM:
     MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
     COMPUTER: IBM PC compatible
     OPERATING SYSTEM: PC-DOS/MS-DOS
```

Search completed: March 9, 2004, 14:01:05
Job time : 167 secs

```
                    GenCore version 5.1.6
            Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:        March  9, 2004, 12:40:58 ; Search time 1291 Seconds
                                           (without alignments)
                                           4833.136 Million cell updates/sec

Title:          US-10-033-109-3
Perfect score:  1706
Sequence:       1 gcacgagcactcccaaccc........aaaaaaaaaaaaaaaa 1706

Scoring table:  IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched:       2421054 seqs, 1828716029 residues

Total number of hits satisfying chosen parameters:      4842108

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :      Published Applications NA:*
              1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
              2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
              3: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
              4: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
              5: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
              6: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
              7: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
              8: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
              9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
             10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
             11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
             12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
             13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
             14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
             15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
             16: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
             17: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
             18: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                              SUMMARIES
```

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1706	100.0	1706	13	US-10-033-109-3	Sequence 3, Appli
2	1666.2	97.7	2162	12	US-10-424-599-139317	Sequence 139917,
3	768.6	45.1	1991	13	US-10-033-109-5	Sequence 5, Appli
4	755.8	44.3	1506	9	US-09-938-842A-1203	Sequence 1203, Ap
5	755.8	44.3	1506	11	US-09-938-842A-1203	Sequence 1203, Ap
6	755.8	44.3	1748	8	US-08-635-967-1	Sequence 1, Appli
7	743.6	43.6	2472	12	US-10-424-599-54229	Sequence 54229, A
8	705.8	41.4	1545	9	US-09-938-842A-1490	Sequence 1490, Ap
9	705.8	41.4	1545	11	US-09-938-842A-1490	Sequence 1490, Ap
10	705.8	41.4	1497	9	US-09-938-842A-1887	Sequence 1887, Ap
11	700.4	41.1	1497	11	US-09-938-842A-1887	Sequence 1887, Ap
12	488.2	28.6	842	12	US-10-424-599-70091	Sequence 70091, A
13	469	27.5	1498	12	US-10-425-114-32362	Sequence 32362, A
14	314	18.4	459	12	US-10-424-599-109084	Sequence 109084,
15	285.8	16.8	423	12	US-10-424-599-139607	Sequence 139607,

16	247.2	14.5	300	12	US-10-424-599-109085	Sequence 109085,
17	232.2	13.6	636	12	US-10-425-114-36312	Sequence 36312, A
18	221	13.0	1037	13	US-10-033-109-1	Sequence 1, Appli
19	157.4	9.2	885	12	US-10-425-114-33461	Sequence 33461, A
20	152.4	8.9	463	12	US-10-424-599-119759	Sequence 119759, A
21	143.2	8.4	395	12	US-10-424-599-84538	Sequence 84538, A
22	142.2	8.3	445	12	US-10-424-599-89931	Sequence 89931, A
23	121	7.1	445	9	US-09-770-444-799	Sequence 799, App
24	109	6.4	197	12	US-10-424-599-89996	Sequence 89996, A
25	97	5.7	9025608	14	US-10-156-761-1	Sequence 1, Appli
26	95.6	5.6	1308	14	US-10-156-761-5576	Sequence 5576, Ap
27	90	5.3	371	12	US-10-424-599-110503	Sequence 110503,
28	86.2	5.1	1338	14	US-10-156-761-2638	Sequence 2638, Ap
29	79.8	4.7	530	13	US-10-101-487-73	Sequence 73, Appl
30	79.2	4.6	522	13	US-10-101-487-71	Sequence 71, Appl
31	79.2	4.6	554	13	US-10-101-487-69	Sequence 69, Appl
32	79.2	4.6	554	13	US-10-101-487-106	Sequence 106, App
33	78.8	4.6	520	14	US-10-184-644-332	Sequence 332, App
34	78.8	4.6	520	14	US-10-184-634-332	Sequence 332, App
35	76.6	4.5	390	9	US-09-790-399-7	Sequence 7, Appli
36	76	4.5	1926	14	US-10-294-804-3	Sequence 3, Appli
37	76	4.5	8705	14	US-10-291-230-14	Sequence 14, Appl
38	76	4.5	8705	14	US-10-291-249-14	Sequence 14, Appl
39	76	4.5	9600	14	US-10-278-751-1	Sequence 1, Appli
40	76	4.5	10233	14	US-10-050-898-283	Sequence 283, App
41	76	4.5	10285	14	US-10-050-902-283	Sequence 283, App
42	74.4	4.4	96602	15	US-10-085-117-61	Sequence 61, Appl
43	74.2	4.3	1248	9	US-09-861-289-7	Sequence 7, Appli
44	74.2	4.3	1248	9	US-09-860-846-7	Sequence 7, Appli
45	74.2	4.3	1248	10	US-09-988-384B-7	Sequence 7, Appli

```
                              ALIGNMENTS

RESULT 1
US-10-033-109-3
; Sequence 3, Application US/10033109
; Publication No. US2002014390A1
; GENERAL INFORMATION:
;  APPLICANT: Allen, Stephen M.
;  APPLICANT: Rafalski, J. Antoni
;  APPLICANT: Sakai, Hajime
;  TITLE OF INVENTION: Nitrogen Transport Metabolism
;  FILE REFERENCE: BB-1210
;  CURRENT APPLICATION NUMBER: US/10/033,109
;  CURRENT FILING DATE: 2001-12-28
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION APPLICATION NUMBER: 09/384,625
;  PRIOR FILING DATE: EARLIER FILING DATE: 1999-08-27
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION APPLICATION NUMBER: 60/098,248
;  PRIOR FILING DATE: EARLIER FILING DATE: 28 August 1998
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: Microsoft Office 97
; SEQ ID NO 3
;  LENGTH: 1706
;  TYPE: DNA
;  ORGANISM: Glycine max
US-10-033-109-3

  Query Match             100.0%;  Score 1706;  DB 13;  Length 1706;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1706;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy          1 GCACGAGCACTCCCAACCCCACCCGTAGTTTCTACCACCTTCAGTCACGGCGTAATACA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GCACGAGCACTCCCAACCCCACCCGTAGTTTCTACCACCTTCAGTCACGGCGTAATACA 60

Qy         61 CTAACCAACCACCACCATGTCGCTGCCTGCTTGTCCCGCGAACAACTGGCCAACTTCTCG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTAACCAACCACCACCATGTCGCTGCCTGCTTGTCCCGCGAACAACTGGCCAACTTCTCG 120

Qy        121 GCCCAAACACCACAGACGCCGCTCCGCGCGCCTCCCTTATCTGCGGCCATTTCGCCGCCG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Qy   121 GCCCAACACCACCAGACGACGCCTCCCGCCGCGCCTATCTGCGGCCCATTTCGCCGCCG 180

Qy   181 TGGACAGCAAGTTCGTCGACACGGCCTTCGCCGTCGACAACACCTACCTCCTCTTTTCCG 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 TGGACAGCAAGTTCGTCGACACGGCCTTCGCCGTCGACAACACCTACCTCCTCTTTTCCG 240

Qy   241 CCTACCTCGTTTTTCTATGCAGCTCGGCTTCGCCATGCTCGCGCCGGCTCCGTCCGCG 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 CCTACCTCGTTTTTCTATGCAGCTCGGCTTCGCCATGCTCGCGCCGGCTCCGTCCGCG 300

Qy   301 CCAAGAACACCATGAACATCATGCTCACCAAGTCCTGGACGCTGCGCCGCGGGGCCTCT 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 CCAAGAACACCATGAACATCATGCTCACCAAGTCCTGGACGCTGCGCCGCGGGGCCTCT 360

Qy   361 TCTACTACCTCTTCGGCTTCGGCCTTTCGGCTCCCCCTTCCAAGGCTTCATCGGTA 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 TCTACTACCTCTTCGGCTTCGGCCTTTCGGCTCCCCCTTCCAAGGCTTCATCGGTA 420

Qy   421 AACATTTCTTCGGCCTCAAGGACATCCTCATCCTCCTAGACTACAGCTACTTCCTCT 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 AACATTTCTTCGGCCTCAAGGACATCCTCATCCTCCTAGACTACAGCTACTTCCTCT 480

Qy   481 ACCAATGGGCCTTCGCCATCGCCGCCGGATCACCAGCCGAAGCATCGCCGAACGCA 540
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 ACCAATGGGCCTTCGCCATCGCCGCCGGATCACCAGCCGAAGCATCGCCGAACGCA 540

Qy   541 CACAGTTCGTGGCCTATCTCATCTCACCGGCTTCGTCACCGGTGG 600
         |||||||||||||||||||||||||||||||||||||||||||||
Db   541 CACAGTTCGTGGCCTATCTCATCTCACCGGCTTCGTCACCGGTGG 600

Qy   601 TCTCCCACTGGTTCGGTCCCCAGACGGCTGGGCCTCTGCCTTTAAGATCACCGACCGGC 660
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 TCTCCCACTGGTTCGGTCCCCAGACGGCTGGGCCTCTGCCTTTAAGATCACCGACCGGC 660

Qy   661 TATTTTCCACCGGCGTAATAGACTTCGCCGGTTCCGGCTAGTCCACATGGTCGGCCGAA 720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 TATTTTCCACCGGCGTAATAGACTTCGCCGGTTCCGGCTAGTCCACATGGTCGGCCGAA 720

Qy   721 TAGCCGGCCTATGGGGAGCGCTGATCGAAGGCCCAAGAATGGGACGTTTCGATCATGCAG 780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 TAGCCGGCCTATGGGGAGCGCTGATCGAAGGCCCAAGAATGGGACGTTTCGATCATGCAG 780

Qy   781 GACGAGCTGTGGCCTTGCGAGGCCACAGCGCGCTCCTTAGTAGTCCTGGGAACCTTCTTGC 840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 GACGAGCTGTGGCCTTGCGAGGCCACAGCGCGCTCCTTAGTAGTCCTGGGAACCTTCTTGC 840

Qy   841 TTTGGTTCGGTTGGTACGGATTTAACCCGGTTCATTTAACAAAATCCTACTTACTTACG 900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 TTTGGTTCGGTTGGTACGGATTTAACCCGGTTCATTTAACAAAATCCTACTTACTTACG 900

Qy   901 GTAACTCAGGAAATTACTACGGTCAATGGAGCCGGTTGGCAGAACCGCGGTCACCACTA 960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 GTAACTCAGGAAATTACTACGGTCAATGGAGCCGGTTGGCAGAACCGCGGTCACCACTA 960

Qy   961 CCCTAGCGGGGGTCAACAGCTGCCTTGACCACGCTATTCGGTAAACGGGTGATATCCGGTC 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 CCCTAGCGGGGGTCAACAGCTGCCTTGACCACGCTATTCGGTAAACGGGTGATATCCGGTC 1020

Qy  1021 ACTGAACGTGACCGATGTCTGCAACGGGCTGTTAGGCGGTTTTGTTGCTTCATAAGTAT 1080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021 ACTGAACGTGACCGATGTCTGCAACGGGCTGTTAGGCGGTTTTGTTGCTTCATAAGTAT 1080

Qy  1081 GTTGCTTCCGTGGTTGAGCCATGGCAGCCATCGTATGCGGTTTCGACGATCCTCTGGAGGCGGCGC 1140
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081 GTTGCTTCCGTGGTTGAGCCATGGCAGCCATCGTATGCGGTTTCGACGATCCTCTGGAGGCGGCGC 1140

Qy  1141 TAATAGCTTGCAACAAATTAGCAGAGAGAAGGTTAAGTTCGACGATCCTCTGGAGGCGGCGC 1200
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141 TAATAGCTTGCAACAAATTAGCAGAGAGAAGGTTAAGTTCGACGATCCTCTGGAGGCGGCGC 1200

Qy  1201 AGTTGCACGGTGGGGTGTGGCACGTGGGGGGTGATATTCACGGCGTTGTTGCCAAAAAGG 1260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201 AGTTGCACGGTGGGGTGTGGCACGTGGGGGGTGATATTCACGGCGTTGTTGCCAAAAAGG 1260
```

```
Qy  1261 AGTATGTGAAGGAGGTTTACGGGTTGGGGAGGGCGCACGGGTTGCTCATGGGGGTGGTG 1320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1261 AGTATGTGAAGGAGGTTTACGGGTTGGGGAGGGCGCACGGGTTGCTCATGGGGGTGGTG 1320

Qy  1321 GGAAGTTGCTGCGGCGCACGTGATTCAGATTCGTGGTGATTGCTGTGGGTGGTTAGTGCGA 1380
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1321 GGAAGTTGCTGCGGCGCACGTGATTCAGATTCGTGGTGATTGCTGTGGGTGGTTAGTGCGA 1380

Qy  1381 CCATGGACCCTGTTGTTTGGGGGTGAATAAACTGAAGCTGTTGAGGATTTCTTCAGAGG 1440
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1381 CCATGGACCCTGTTGTTTGGGGGTGAATAAACTGAAGCTGTTGAGGATTTCTTCAGAGG 1440

Qy  1441 ATGAGCTTGCGGGGATGGACATGACTCGCCATGGAGGGCTTGCTTATGCTTATGAGGATG 1500
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1441 ATGAGCTTGCGGGGATGGACATGACTCGCCATGGAGGGCTTGCTTATGCTTATGAGGATG 1500

Qy  1501 ATGAGACGCACAAGCACATGGGATGCAGTGAGGAGGGGTTGGGCCCAACGGCTCTTCCACAC 1560
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1501 ATGAGACGCACAAGCACATGGGATGCAGTGAGGAGGGGTTGGGCCCAACGGCTCTTCCACAC 1560

Qy  1561 CCACCACTGATGATGAATGATCTTTTTCCCATATGCATGTCCATATGAGTCATTAGTCAAACATTAAAT 1620
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1561 CCACCACTGATGATGAATGATCTTTTTCCCATATGCATGTCCATATGAGTCATTAGTCAAACATTAAAT 1620

Qy  1621 TTGGATACATATTCCTTGTAAGGATTCAAACCTTGGTTACTTGTTACTTCTGTTAAAAAA 1680
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1621 TTGGATACATATTCCTTGTAAGGATTCAAACCTTGGTTACTTGTTACTTCTGTTAAAAAA 1680

Qy  1681 AAAAAAAAAAAAAAAAAAAAAAAAAAAA 1706
         |||||||||||||||||||||||||||
Db  1681 AAAAAAAAAAAAAAAAAAAAAAAAAAAA 1706
```

```
RESULT 2
US-10-424-599-139317
; Sequence 139317, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa Thomas J
;  APPLICANT: Kovalic David K
;  APPLICANT: Zhou Yihua
;  APPLICANT: Cao Yongwei
;  TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53223)B
;  CURRENT APPLICATION NUMBER: US/10/424,599
;  CURRENT FILING DATE: 2003-04-28
;  NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 139317
;  LENGTH: 2162
;  TYPE: DNA
;  ORGANISM: Glycine max
;  FEATURE:
;  OTHER INFORMATION: Clone ID: PAT_MRT3847_96810C.1
US-10-424-599-139317

  Query Match          97.7%;   Score 1666.2;   DB 12;   Length 2162;
  Best Local Similarity  99.5%;   Pred. No. 0;
  Matches 1671;  Conservative  0;  Mismatches  8;  Indels  0;  Gaps  0;

Qy    4 CGAGCACTCCCAACCCCACCCGTAGTTCTACCACCTTCAGTCACGGCGTAATACACTA 63
        ||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db   47 CCACCACTCCCAACCCCACCCACCCGTAGTTTCTACCACCTTCAGTCACGGCGTAATACACTA 106

Qy   64 ACCAACCCACCATGTCCGCTGCCTGCTTGTCCCGCGGACAACTGGCCCAACTTCTCGGCC 123
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  107 ACCAACCCACCATGTCCGCTGCCTGCTTGTCCCGCGGACAACTGGCCCAACTTCTCGGCC 166

Qy  124 CAAACACCACAGACGCCTCCGCGCCGCCTCCCTTATCTGCGGCCATTTCGCCGCCGTGG 183
        |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db  167 CAAACACCACAGACGCCTCCGCGCCGCCTCCCTTATCTGCGGCCAGTTCGCCGCCGTGG 226
```

```
Qy   184  ACAGCAAGTTCGTCGACACGGCCTTCGCCGTCGACAACACCTACCTCCTCTTTCCGCCT  243
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   227  ACAGCAAGTTCGTCGACACGGCCTTCGCCGTCGACAACACCTACCTCCTCTTTCCGCCT  286

Qy   244  ACCTCGTTTTTCTATGCAGCTCGGCTTCGCCATGTCTGCGGCCGCTCCGTCCGCGGCCA  303
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   287  ACCTCGTTTTTCTATGCAGCTCGGCTTCGCCATGTCTGCGGCCGCTCCGTCCGCGGCCA  346

Qy   304  AGAACACCATGAACATCATGCTCACCAACGTCCTGAGGCTGCGGCCGGCCTCTTCT     363
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   347  AGAACACCATGAACATCATGCTCACCAACGTCCTGACGCTGCGGCCGGCCTCTTCT     406

Qy   364  ACTACCTCTTCGGCCTTCGCCTTCGGCGTCCCCCTCCAACGGCTTCATCGGTAAAC    423
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   407  ACTACCTCTTCGGCCTTCGCCTTCGGCGTCCCCCTCCAACGGCTTCATCGGTAAAC    466

Qy   424  ATTTCTTCGGCCTCAAGGACATCCCTTCATCCTCTACGACTACAGCTACTTCCTCTACC  483
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   467  ATTTCTTCGGCCTCAAGGACATCCCTTCATCCTCTACGACTACAGCTACTTCCTCTACC  526

Qy   484  AATGGGCCTTCGCCATCGCCCGGCATCACCACCAGCGAAGCATCGCCGAACGCACAC    543
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   527  AATGGGCCTTCGCCATCGCCCGGCATCACCACCAGCGAAGCATCGCCGAACGCACAC    586

Qy   544  AGTTCGTGGCCTATCTCATCTACTCCTCTTCCTCACCGGCTTCGTCTATCCGTGGTCT  603
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   587  AGTTCGTGCGCCTATCCATCTACTCCTCTTCCTCACCGGCTTCGTCTATCCGTGGTCT  646

Qy   604  CCCACTCGGTTCTCGGTCCCCAGACGACGGCTGGGACCCTGCCTTTAAGATCACCGACCGGCTAT  663
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   647  CCCACTCGGTTCTGGGTCCCCAGACGACGGCTGGGGCCTCTGCCTTTAAGATCACCGACCGGCTAT  706

Qy   664  TTTCACCCGGCGTAATAGACTTCGGCCGCGGTTCCGGCCGTAGTCCACATGGTCGGCGGAATAG  723
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   707  TTTCCACCCGGCGTAATAGACTTCGGCCGCGGTTCCGGCCGTAGTCCACATGGTCGGCGGAATAG  766

Qy   724  CCGGCCTATGGGGAGGCGCTGATCGAAGGGCCCAAGAGATGGACGTTTCGATCATGCAGGAC  783
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   767  CCGGCCTATGGGGAGGCGCTGATCGAAGGGCCCAAGAGATGGGACGTTTCGATCATGCAGGAC  826

Qy   784  GAGCTGTGCCTTGCGAGGGCCACAGCCGCGTCCTTAGTAGTCCTGGGAACCTTCTTGCTTT  843
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   827  GAGCTGTGCCTTGCGAGGGCCACAGCCGCGTCCTTAGTAGTCCTGGGAACCTTCTTGCTTT  886

Qy   844  GGTTCGGTTGGTACGGATTAACCCCGGTTCATTTAACAAAATCCTACTTACTTACGGTA    903
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   887  GGTTCGGTTGGTACGGATTTAACCCCGGTTCATTTAACAAAATCCTACTTACTTACGGTA    946

Qy   904  ACTCAGGAAATTACTACGGTCAATGAGCGCGGGTTGGCAGAACCGCGGTCACCACTACCC   963
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   947  ACTCAGGAAATTACTACGGTCAATGGAGCGCGGGTTGGCAGAACCGCGGTCACCACTACCC  1006

Qy   964  TAGCGGGGTCAACAGCTGCCTTGACCACGCTATTCGGTGAAATCCGGTCACT          1023
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1007  TAGCGGGGTCAACAGCTGCCTTGACCACGCTATTCGGTGAAATCCGGTCACT          1066

Qy  1024  GGAACGTGACCGATGTCTGCAACGGGCTGTTAGGCGGTTTCGCGGCGGATAACAGCCGGTT  1083
          |||||||||||||||||||||||||||| ||||||| |||| ||||||||||||||||| |
Db  1067  GGACGTGACCGATGTCTGCAACAGGTTCGCCGGACACCACGTCCACTTACCTGCTC         1126

Qy  1084  GCTCCGTGGTTGAGCAAGGGCAGCCAATCGTATGCGGGTTTGTTGCTTCTATAGTATTAA  1143
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1127  GCTCCGTGGTTGAGCAAGGGCAGCCAATCGTATGCGGGTTTGTTGCTTCTATAGTATTAA  1186

Qy  1144  TAGCTTGCAACAAATTAGCAGAGAAGGTTAAGTTCGACGATCCTCTGGAGGCGGCGCAGT  1203
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1187  TAGCTTGCAACAAATTAGCAGAGAAGGTTAAGTTCGACGATCCTCTGGAGGCGGCGCCAGT  1246

Qy  1204  TGCACGGTGGGGTGTGGCACGTGGGGGGGTGATATTCGCAAAAAGGAGT  1263
          ||||||||||||||||||||||||||||||||||||||||||| |||||
Db  1247  TGCACGGTGGGGTGTGGCACGTGGGGGGGTGATATTCGCAAAAAGGAGT  1306

Qy  1264  ATGTGAACGAGAGGTTTACGGGTTGGGAGGGCGCACGGTGGGTGTGGGAG  1323
```

```
Db  1307  ATGTGAAGGAGAGGTTTACGGGTTGGGGAGGGGCGCACGGGTGGTGTGGGA  1366

Qy  1324  AGTTGCTGCGGCCGCGCACGTGATTCAGATTCGTGATTGCTGGTGGGTTAGTGCGACCA  1383
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1367  AGTTGCTGCGGCCGCGCACGTGATTCGGTGATTCGGTGATTGCTGGGTGGTGTTAGTGCGGACCA  1426

Qy  1384  TGGGACCCTTGCTTGTTTGGGGTGAATAAACTGAAGCTGTTGAGGATTCTTCAGAGGATG  1443
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1427  TGGGACCCTTGCTTGTTTGGGGTGAATAAACTGAAGCTGTTGAGGATTCTTCAGAGGATG  1486

Qy  1444  AGCTTGCGGGATGGACATGACTCGCCATGAGGAGGCTTTGCTTATGCTTATGAGAGATGATG  1503
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1487  AGCTTGCGGGATGGACATGACTCGCCATGAGGAGGCTTTGCTTATGCTTATGAGAGATGATG  1546

Qy  1504  AGACGGACAAGCATGGATGGCAGTTGAGGAGGGTTGGGCCCAACGCGTCTTCCACCCA  1563
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1547  AGACGGACAAGCATGGATGGCAGTTGAGGAGGGTTGGGCCCAACGCGTCTTCCACACCCA  1606

Qy  1564  CCACTGATGAATGATCTTTTTTCCCATATGCATGTCTCATTAGTCAAACATTAAATTTG  1623
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1607  CCACTGATGAATGATCTTTTTTTTCCCATATGCATGTCTCATTAGTCAAACATTAAATTTG  1666

Qy  1624  GATACATATTCCTTGTAAGGATTCAAACCTTGGTTACTTGTTACTTCTGTTAAAAAAAA  1682
          |||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db  1667  GATACATATTCCTTGTAAGGATTCAAACCTTGGTTACTTGTTACTTCTGTTAGATCCAA  1725
```

```
RESULT 3
US-10-033-109-5
; Sequence 5, Application US/10033109
; Publication No. US20020142390A1
; GENERAL INFORMATION:
;   APPLICANT: Allen, Stephen M.
;   APPLICANT: Rafalski, J. Antoni
;   APPLICANT: Sakai, Hajime
;   TITLE OF INVENTION: Nitrogen Transport Metabolism
;   FILE REFERENCE: BB-1210
;   CURRENT APPLICATION NUMBER: US/10/033,109
;   CURRENT FILING DATE:     2001-12-28
;   PRIOR APPLICATION NUMBER: EARLIER APPLICATION APPLICATION NUMBER: 09/384,625
;   PRIOR FILING DATE: EARLIER FILING DATE: 1999-08-27
;   PRIOR APPLICATION NUMBER: EARLIER APPLICATION APPLICATION NUMBER: 60/098,248
;   PRIOR FILING DATE: EARLIER FILING DATE: 28 August 1998
;   NUMBER OF SEQ ID NOS: 14
;   SOFTWARE: Microsoft Office 97
; SEQ ID NO 5
;   LENGTH: 1991
;   TYPE: DNA
;   ORGANISM: Triticum aestivum
US-10-033-109-5

  Query Match             45.1%;  Score 768.6;  DB 13;  Length 1991;
  Best Local Similarity   73.8%;  Pred. No. 2.2e-193;
  Matches  991;  Conservative    0;  Mismatches  349;  Indels    3;  Gaps    1;

Qy   174  GCCGCCGTGGACAGCAGCAAGTTCGTCGACACGGCCTTGCGCCGTCGACAACACCTACCTCCTC  233
          ||||||||||||||||| ||||||||||||| ||||| |||||||||||||||||||||
Db   134  GACTACCTGTGCAACAGGTTCGCCGGACACCACGTCCACTTACCTCCACTTACCTGCTC  193

Qy   234  TTTTCCGCCTACCTCGTTTTTCTATGCAGCTCGGCTTCGCCATGTCTGCGGCCGCTCC  293
          |||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   194  TTTCTCGGCCTACCTCGTCTCTCGCCATGCAGCTCGGCTTCGCCATGTCTGCGGCCGCTCC  253

Qy   294  GTCCGCGGCCAAGAACACCATGAACATCATGCTCACCAACGTCCTGAGGCTGCGGCCGGC  353
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db   254  GTCCGGGCCAAGAACAACACCATGAACATCATGCTCACCAACGTCCTGACGCTGCGGCCGGC  313

Qy   354  GGCTCTTCTACTACCTCTTCGGCCTTCGGCCTTCGGCGTCCCCCTCCAACGGCTTC  413
          ||||||||||| ||||||||||||||||||||||||||||| |||| |||||||| 
Db   314  GCGCTCTTCTACTACCTCTTCGGCCTTCGGCCTTCGGACGCCTCGGACGGCGTTC  373

Qy   414  ATCGGTAAACATTTCTTCGGCCTCAAGGACATCCCTTCATCCTCTACGACTACAGCTAC  473
```

Db   1454 TACCACGACGACGACGAGCACGA 1476

RESULT 4
US-09-938-842A-1203
; Sequence 1203, Application US/09938842A
; Patent No. US2002160378A1
; GENERAL INFORMATION:
;   APPLICANT: Harper, Jeff
;   APPLICANT: Kreps, Joel
;   APPLICANT: Wang, Xun
;   APPLICANT: Zhu, Tong
;   TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
;   TITLE OF INVENTION: SAME, AND METHODS OF USE
;   FILE REFERENCE: SCRIP1300-3
;   CURRENT APPLICATION NUMBER: US/09/938,842A
;   CURRENT FILING DATE: 2001-08-24
;   PRIOR APPLICATION NUMBER: US 60/227,866
;   PRIOR FILING DATE: 2000-08-24
;   PRIOR APPLICATION NUMBER: US 60/264,647
;   PRIOR FILING DATE: 2001-01-16
;   PRIOR APPLICATION NUMBER: US 60/300,111
;   PRIOR FILING DATE: 2001-06-22
;   NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1203
;   LENGTH: 1506
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
US-09-938-842A-1203

Query Match   44.3%;   Score 755.8; DB 9; Length 1506;
Best Local Similarity   71.7%;   Pred. No. 4.8e-190;
Matches 1036; Conservative   0; Mismatches 397; Indels   12; Gaps   3;

Qy   136 ACGGCTCCGCGCCGCCTCCCTTATCTGCGGCCATTTCGCCGCCGTGGACAGCAAGTTCG 195
           |||| |||||| |||||| || |||||||||||||||||||||||||| |||||||||
Db   44 ATGCCACGGCGGCGCCAACTCACATCTGTGGCAGTTAGGCGACGTCAACAACAAATTTA 103

Qy   196 TCGACACGGCCTTCGCCTGCGTCGACAACACCTACCTCCTTTTCGCCTCGTTTTTT 255
           ||||| || ||||||||||||||| ||| ||||||||| ||||||||||||||
Db   104 TCGACACCGCTTTCGCTATAGACAACACACTTACCTTCTCTCCGCCTACCTTGTCTTCT 163

Qy   256 CTATGCAGCTCGGCTTCGCCATGCTCTGCGCCGGCTCCGTCCGCCAAGAACACCATGA 315
           |||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db   164 CTATGCAGCTTGGCTTCGCTATGCTCTGCGTGCCGGTTCCGTGAGGAGCCAAGAATACTATGA 223

Qy   316 ACATCATGCTCACCAACGTCCTGGACGCTGCGACGCTGCTCTTCTACTACCTCCTTCG 375
           ||||||||||||||||| |||||||||||||||||||| |||||||||||||||||||
Db   224 ACATCATGCTTACCAACGTCCTTGACGCTGCAGCGCCGGTGGTCTTCTTCTATTATCTGTTTG 283

Qy   376 GCTTCGGCCTTCGCTTTCGGCTCCCCCTCCAACGCGTCATCGTCATCCGTAAACATTTCTTCGGCC 435
           ||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db   284 GCTACGGCCTTTGCCTTTGGCGATCTCCGTCCAATGGTTTCATCGGTAAACACTACTTTGGTC 343

Qy   436 TCAAGGACATCCCTTCATCCTCCTACGACTACAGCTACTTCCTCTACCAATGGCCTTCG 495
           ||||||||||| ||||||||| |||||||| ||||||||||| |||||||||||| |||
Db   344 TCAAAGACATCCCCACGGCCTTCGCTGACTACTCCCAACTTTCTCTACCAATGGGCCTTTTG 403

Qy   496 CCATCGCGCCGCCGGCATCACCAGCGGAAGCATCGCCGAACGACACCAGTTCGTGGCCT 555
           |||||||||||||||||||||||| |||| |||| |||||||| ||||| |||||| ||
Db   404 CAATCGCTGCGGCGTGGAATCACAAGTGGCTCGATCGGTCGATCGACCAGTTCGTGGGCTT 463

Qy   556 ATCTCATCTACTCCTTCCTCACCGGCTTCGTCTATCCGGTGGTCTCCCACTGGTTCT 615
           |||||||| ||||| |||||||| |||||||| ||||||||||||||||||||||||
Db   464 ACCTAATCTATTCCTCTTTCTTAACCGGTTTGTTTACCCGGTCGTCTCCTCACTGGTTCT 523

Qy   616 GGTCCCCAGAGCGGCTGGGCCTTCGGCCTTGCCTT----TAAGATCACCGACGCGGCTATTTTCCACCG 672
           ||||||||||||||||||||| |||||||||| |||||||||||||||||||||||
Db   524 GGTCAGTTGATGGATGGGCCAGCCCGTTCCGTACCGATGGAGAATTGCTTTTCAGCACCG 583

Qy   673 GCGTAATAGACTTCGCCGGTTCGGCGTAGTCCACATGGTCGGCGGAATAGCCGGCCTAT 732
           || ||||||| |||||||| |||||||||||||||||||||||||||||||||||||
Db   584 GAGCGATAGATTTCGCTGGCGTGGTGTTGTTCATATGGTCGGAGGTATCGCGGGGTGACTCT 643

```
;    FILE REFERENCE: SCRIP1300-3
;    CURRENT APPLICATION NUMBER: US/09/938,842A
;    CURRENT FILING DATE: 2001-08-24
;    PRIOR APPLICATION NUMBER: US 60/227,866
;    PRIOR FILING DATE: 2000-08-24
;    PRIOR APPLICATION NUMBER: US 60/264,647
;    PRIOR FILING DATE: 2001-01-16
;    PRIOR APPLICATION NUMBER: US 60/300,111
;    PRIOR FILING DATE: 2001-06-22
;    NUMBER OF SEQ ID NOS: 5379
;  SEQ ID NO 1203
;    LENGTH: 1506
;    TYPE: DNA
;    ORGANISM: Arabidopsis thaliana
US-09-938-842A-1203

Query Match          44.3%;   Score 755.8;  DB 11;  Length 1506;
Best Local Similarity  71.7%;  Pred. No. 4.8e-190;
Matches 1036;  Conservative  0;  Mismatches 397;  Indels  12;  Gaps  3;

Qy   136 ACGCCTCCGCCGCCGCCTTCCCTTATCTGCGGCCATTCGCCGCCGTGGACAGCAAGTTCG 195
                          |||| ||||||||| |||| |||||||||||||
Db    44 ATGCCACGGCGCGGCCCAACTACATCTGTGGCCAGTTAGGCGACGTCAACAACAAATTTA 103

Qy   196 TCGACACGGCCTTCGCCGTCGACACACCTACCTCCTCCTTTTCGGCCTACCTCGTTTTTT 255
             ||||||||||||||||| ||||| ||  |  |  | |||||||||||||||||||
Db   104 TCGACACCGCCTTTCGCTATAGACAACACTTACCTTCTCTCCGCCTACCTTGTCTTCT 163

Qy   256 CTATGCAGCTCGGCTTCGCCATGCTCGTGCCGCCGGCCTCCGTGCCGCCAAGAACACCATGA 315
             ||||||||||||| |||||| ||||||||||||| ||||| |||||||
Db   164 CTATGCAGCTGGCTTGCTTGCTATGCTCGTGTGCCGGTTCCGGTGCCGGTTCCGGAGAGCCAAGAATACTATGA 223

Qy   316 ACATCATGCTCACCAACGTCCTGGACGCTGCCGCCGGCGCCTCTTCTACTACCTCCTTCG 375
             ||||||||| ||||||||| |||| ||||| ||  ||||| |||| |||||||||||
Db   224 ACATCATGTTACCAACGTCCTTGACGCTGCAGCGCTGCGTGTCTCTTCTATTATCTGTTTG 283

Qy   376 GCTTCGCCTTCGCTTCGGCTGCCCCCTCCAACGCTGCCTCATCGGTAAACATTTCTTCGGCC 435
             |||||||||||  |||  ||||| |||||||||| ||||||| ||||||| ||||||
Db   284 GCTACGCCTTTGCTTTGGATCTCCGTCCAATGTTTCATCGGTAAACACTACTTTGGTC 343

Qy   436 TCAAGGACATCCTTCATCCTCCTACGACTACAGCTACTTCCTCTACCAATGGGCCTTCG 495
             |||||||||||| ||| |||||||||||||||||||||||||||  |||||||||||||
Db   344 TCAAAGACATCCCCACGGCCTCTGCTGCTACTCCAACTTCTCTACCAAGTGGGCCTTTG 403

Qy   496 CCATCGCCGCGCCGCCGGCATCACCAGCCGAAGCATCGCCGAACGACACACAGTTCGTGGCCT 555
             |||||| ||| || ||||| ||||||||||| ||||||||||||| || ||||||||||
Db   404 CAATCGCTGCGCGGCTGGAATCACAAGTGCCTGCTGCTCGATGCTCGATGCTGCTCGTGGCTT 463

Qy   556 ATCTCATCTACTCCTCCTTCCTCCACCGGCTTCGTCGTCGCTCCGGTTGTGGTCTCCCACTGGTTCT 615
             |||||||||||| ||| ||| |||||||| |||| |||||| |||||||||||
Db   464 ACCTAATCTACTTCCTCTTTCTTTAACCGGGTTGTTGTTCACCGGTCGTCTCACTGGTTCT 523

Qy   616 GGTCCCCAGACGGCTGGGCCCTCTGCCT---TAAGATCACCGACCGGCTATTTTCCACCG 672
             |||| |||||| ||| |||||||||||||      |||| || || ||||| || |||||
Db   524 GGTCAGTTGATGATGGCCAGCCCGTTCCGGTACCGGTACCGATGGAGATTTGCTTTTCAGCACCG 583

Qy   673 GGTAATAGACTTCGCCGGTTCCGGCGTAGTCCACATGTCGGCGGCGGAATAGCCGGCCTAT 732
             ||||| ||||||| ||||||||| |||| ||||||||||| |||||||| ||||| |||||
Db   584 GAGCCGATAGATTTCGCTGGGTCCGGTGTTGTTCATATGTCGGAGGTATCGGCGGATCTCT 643

Qy   733 GGGGAGCGCTGATCGATGGAAGGGCCCAAGAATGGACGTTTCGATCATGCAGGACGAGCTGTGG 792
             |||||||||| ||| | ||||| ||||||||| ||| |||  |||||||||||||||||||||
Db   644 GGGGTGCGCTCATCGAAGGTCCACGACTTGGCCGGTTCGATAACGGAGGCCGTGCCCATCG 703

Qy   793 CCTTGGAGGCCACAGCGCGTCCTTAGTCGTTCCTGGGAACCTTCTTCGTTGGTTCGGTT 852
             |||| ||||||||||| |||||||| ||||| |||| ||||| ||||| ||||||||||
Db   704 CTCTTCGTGGCCACTCGCCGTCACTTGTTGTCCTTCGTTGGAACATTCCTCCCTCTGGTTTGGAT 763

Qy   853 GGTACGGCGATTTAAACCCCGGTCATTTAACAAAATCCTACTTACTTACGGTAACTCAGGAA 912
             |||||||||||| |||||||||||| |||| |||| ||||||||||| |||||||||||
Db   764 GGTACGGATTTAACCCCGGTCCTTCAACAGATCCTAGTCACGTACGAGA---CAGGCA 820

Qy   913 ATTACTACGGTCAATGAGCCGCGGTTGGCAGAACCGGCGCGTTCACCACTACCCTAGCGGGGT 972
```

```
FILING DATE: 29-APR-1996
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
   APPLICATION NUMBER: PCT/EP 94/03499
   FILING DATE: 24-OCT-1994
PRIOR APPLICATION DATA:
   APPLICATION NUMBER: DE P 43 37 597.9
   FILING DATE: 28-OCT-1993
ATTORNEY/AGENT INFORMATION:
   NAME: Meilman, Edward
   REGISTRATION NUMBER: 24,735
   REFERENCE/DOCKET NUMBER: P/951-120
TELECOMMUNICATION INFORMATION:
   TELEPHONE: (212) 382-0700
   TELEFAX: (212) 382-0888
   TELEX: 236925
INFORMATION FOR SEQ ID NO: 1:
   SEQUENCE CHARACTERISTICS:
      LENGTH: 1748 base pairs
      TYPE: nucleic acid
      STRANDEDNESS: single
      TOPOLOGY: linear
   MOLECULE TYPE: cDNA
   ORIGINAL SOURCE:
      ORGANISM: Arabidopsis thaliana
      INDIVIDUAL ISOLATE: Ammonium transporter
   IMMEDIATE SOURCE:
      LIBRARY: cDNA library in plasmid pF161
   POSITION IN GENOME:
      MAP POSITION: from 21 to 1523 coding region
   FEATURE:
      NAME/KEY: CDS
      LOCATION: 21..1526
US-08-635-967-1

Query Match       44.3%;        Score 755.8;  DB 8;  Length 1748;
Best Local Similarity 71.7%;    Pred. No. 5.2e-190;
Matches 1036;  Conservative  0;  Mismatches 397;  Indels 12;  Gaps 3;

Qy  136 ACGCTCCGCGCCGCCTCCGCCTCCCTTATCTGCGGCCATTCGCCGCCGTGGACAGCAAGTTCG 195
                                                    | | |||  | || || |||||
Db   64 ATGCCACGGCGCGGCGCCAACTACATATGTGGCCAGCGACGTCAACACAAATTCA        123

Qy  196 TCGACAACGGCCTTCGGCGTCGACAACACCTACCTCCTCTTTCCGCCTACCTCGTTTTTT   255
        ||||| ||||| || |  ||| ||| ||| || |||||| | || ||||| ||| | ||
Db  124 TCGACACCGGTTTCGCTATAGACAACACTTACCTCCTCTCCGCCTACCTTGTCTTCT      183

Qy  256 CTATGCAGCTCGGCTTCGCCATGCTCGCGCGGCGTCCGCGCCCAAGAACACCATGA       315
        |||||||||| |||||||| ||||||| || ||| ||| ||| |||||| | ||||
Db  184 CTATGCAGCTTGGCTTCGCTATGCTCGTGTGCCGGTTCCGTGCGGAGAGCCAAGAATACTATGA 243

Qy  316 ACATCATGCTCACCAACGTCCTGGACGCTGCGCGGCGCCTCTTCTACTACCTCTTCG     375
        ||||||||| ||||||||| || |||||| |  | |   | || |||| ||| ||||
Db  244 ACATCATGCTTACCAACGTCCTTGACGCTGCCAGCTAGGCGAGCGTCAACAACAATTCA    303

Qy  376 GCTTCGCCTTTCGCTTCGGCTTCCCCTCCAACGGCTTCATCGGTAAACATTCTTCGGCC    435
        ||   |||||||||| || |||  |   || |||||| |||   |||||  || || ||
Db  304 GCTACGGCCTTTGCTTTCGCTTCCGGTTTCATCGGATCCGGTCCAATGGTTTCATCGGTAAAGACTACTTTGGTC 363

Qy  436 TCAAGGACATCCCTTCATCCTCCTAGCGACTACAGCTACTTCCTCCTACCAATGGGCCTTCG  495
        | ||| ||| ||||   ||   |||| | || ||| |||| | || |||| || |||| |||
Db  364 TCAAAGACATCCCCACGGCCTCTGCTGACTACTCCAACTTTCTACCACCAAATGGGCCTTTG   423

Qy  496 CCATCGCCGCGCCGCCGCCGCATCACCAGCCGGAAGCATCGCCGAACGCACACAGTGTTCGTGGCCT 555
        ||| ||  || || | || |  | ||  |   ||| | ||  ||||  | | | |||| | |||
Db  424 CAATCGCTGCGGCTGAATACAAGGTGCTCGATGCGGTCGTGTTACCGGGTGGCTTCGTGGCTT     483

Qy  556 ATCTCATCTACTCCTCCTTCCTCCACCGGCTTCGTCATCCGGGTGGTCTCCCACTGGTTCT  615
        |||| |  | | |||| || || ||  |    |||  | |||    | ||| || | |||
Db  484 ACCTAATCTATTCCTTCTTCTTTAACCCGGTTTGTTGTTAACACTATTCACTACTTTGGTTCT  543

Qy  616 GGTCCCCAGAGACGGCTGGGCCTCGCCTT---TAAGATCACCAGCGACCGGCTATTTCCACCG  672
        ||| |   |||| | | || |  |  || |||   ||  | ||   |  ||   ||| |||||
```

```
Qy  973 CAACAGCTGCCTTGACCACGCTATTCGGTAAACGGGTGATATCCGGTCACTGGAACGTGA  1032
        |||||| |||||||| || ||| | || ||| ||| | | |||  | || ||||||||| |
Db  881 GCACCGCGGCGGCTGACAACCCTATTTGGGAAACGTCTACTCTCGGGACATTGGAACGGTCA    940

Qy 1033 CCGATGTCTGCAACGGCGTCGTTAGGCGGGTTCGCGGCGGATAACAGCCGCGGTTGCTCCGTGG 1092
        | |||||  |||||||| || || ||||| ||| | || ||| ||| |  | || |||||  ||
Db  941 CTGATGTATGCAACGGCCTCCTCGAGGGTTTGCAGCCATAACTGGTGCGTGCTCTGTCG      1000

Qy 1093 TTGAGCCATGGCAGCCATCGTATGGGTTTGTTTGTTCTTCTATAGTATTAAATAGCTTGCA  1152
        |||||||||||  | || ||| ||||| |  | |||  | |||   | ||| | |  ||| |
Db 1001 TTGAGCCATGGGCTGCGATCATCTGCGGGTTCGTGCGGCGCCCTAGTCCTCCTCGGATGCA   1060

Qy 1153 ACAAATTAGCAGAGAAGGTTAAGTTCGACGATCCTCTCGGAGGCGGCGCAGTTGCACGGTG  1212
        ||||  | |||| ||  | |||| | |||||| | |||  || | | | | |  |||||| |
Db 1061 ACAAGCTCGTGAGAGCTCAAATAGACGACCCCTCTTGAGCAGCACAACTACACCGGTG      1120

Qy 1213 GGTGTGGCACGTGGGGGGGTGATATTCACGGCGTTGTTCGCAAAAAAGGAGTATGTGAAGG  1272
        | || || |  ||||||   |  |||||||| | | | ||| || |  | | || | |
Db 1121 GTTGCGGTGCGTGGGGGGGACTAATATTCACGGCTCTCTTCGCTCAAGAAAAGTACTTGAACC 1180

Qy 1273 AGGTTACGGGTTG-----GGGAGGGCGCACGGGTTGCTCATGGGGGGTGTGGGAAGT  1326
        ||  ||| |     |||||  | ||| | | ||| | ||  |||   || | ||| |
Db 1181 AGATTTACGGCAACAAACCCGGAAGGCCACACACGGTTTGTTTGGCGGTGAGGAGGAAAAC 1240

Qy 1327 TGCTGGCGGCGCGCACGTGATTCGAGATTCTGGTGATTGCTGGGTGGTTAGTGCCGACCATGG 1386
        |  ||| |||| |   |  || | ||||| ||  || | ||| | || | ||||| |||||||
Db 1241 TACTTGGAGTCAGCTAGTCATTCAGATCATTGGATCACGGGTTGGGTAAGTGCGACCATGG    1300

Qy 1387 GACCCTTGTTTGGGGGTGAATAAACTGAAGCTGTTGAGGATTCTTCAGAGGATGAGC   1446
        | || |||| |  | | ||   ||||||    || || |||  || ||| |||||| |
Db 1301 GGACACTTTTTTCTTCATCCTCCAAGAAAATGAAATTGTTGCGGATATCGTCCGAGGATGAGA 1360

Qy 1447 TTGCGGGATGGACATGACTCGCCATGGAGGCTTTGCTTATGCTTATGAGGATGATGAGA  1506
        | ||| | ||||| |||||  | | ||  || |||| || ||||| | | ||||| ||
Db 1361 TGGCCGGTATGGATATGACCAGGCACGGTGGTTTCGTTTGCTTATATATGTACTTTGATGATGATG 1420

Qy 1507 CGCACAAGCACATGGGATGCAGTTGAGGAGGGTTGGGCCCCAACGCGCGTCTTCCACACCCACCA 1566
        | || ||| |  | | | | || |||||| |  || |  | ||  |  ||| || |||  ||
Db 1421 AGTCTCACAAAGCCATTCAGCTTAGGAGAGTTGAGCGTTAGGAGGCACGATCTCCTTTCTCCTTCTGGTG 1480

Qy 1567 CTGAT 1571
        || ||
Db 1481 CTAAT 1485
```

RESULT 6
US-08-635-967-1
Sequence 1, Application US/08635967
Publication No. US20010003848A1

```
GENERAL INFORMATION:
   APPLICANT: FROMMER, Wolf-Bernd
   APPLICANT: NINNEMANN, Olaf
   TITLE OF INVENTION: DNA SEQUENCES FOR AMMONIUM TRANSPORTER,
   TITLE OF INVENTION: PLASMIDS, BACTERIA, YEASTS, PLANT CELLS AND PLANTS
   TITLE OF INVENTION: CONTAINING THE TRANSPORTER
   NUMBER OF SEQUENCES: 2
   CORRESPONDENCE ADDRESS:
      ADDRESSEE: Ostrolenk, Faber, Gerb & Soffen, LLP
      STREET: 1180 Avenue of the Americas
      CITY: New York
      STATE: NY
      COUNTRY: US
      ZIP: 10036-8403
   COMPUTER READABLE FORM:
      MEDIUM TYPE: Floppy disk
      COMPUTER: IBM PC compatible
      OPERATING SYSTEM: PC-DOS/MS-DOS
      SOFTWARE: PatentIn Release #1.0, Version #1.30
   CURRENT APPLICATION DATA:
      APPLICATION NUMBER: US/08/635,967
```

```
;  APPLICANT: Kovalic David K
;  APPLICANT: Zhou Yihua
;  APPLICANT: Cao Yongwei
;  TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53223)B
;  CURRENT APPLICATION NUMBER: US/10/424,599
;  CURRENT FILING DATE: 2003-04-28
;  NUMBER OF SEQ ID NOS: 285684
;  SEQ ID NO 54229
;  LENGTH: 2472
;  TYPE: DNA
;  ORGANISM: Glycine max
;  FEATURE:
;  OTHER INFORMATION: Clone ID: PAT_MRT3847_19982C.1
US-10-424-599-54229

Query Match             43.6%;  Score 743.6;  DB 12;  Length 2472;
Best Local Similarity   73.2%;  Pred. No. 1.1e-186;
Matches 1012;  Conservative  0;  Mismatches 354;  Indels 16;  Gaps  4;

Qy         136 ACGCTCCGCGCCGCCTCCCTTATCGTGGCCATTCGCCGCGTGGACAGCAAGTTCG 195
               ||||||| |||   |||||| ||||| |||||||| |||| |||
Db         585 ACGACACCCGCGCCGCCAACTACCTCTGCGCCCAATTCGATTCCATTTCTAGAAAGCTCA 644

Qy         196 TCGACACGGCCTTCGCCGTCGACAACACCTACCTCCTCTTTTCGCCTACCTCGTTTTTT 255
               ||| |||  |||  || |||||  |||| ||| || ||||| ||||||||||||||||
Db         645 ACGTAACACCTACGGCCGTCGACAACACCTACTTCGTCGTTTTTCAGCGTATCTTGTCTTCG 704

Qy         256 CTATGCAGCTCGGCTTCGCCATGCTTCTGCGCGCGGCTCTGCGCCGCCAAGAACACCATGA 315
               |||||||||| ||||||||| ||| || |||||| |||||||| ||||||||||| |||||
Db         705 CCATGCAGCTCGGCTTCGCCATGCTTCTGCGCGCGGCTCTGCGCCGCCAAAAACACCATGA 764

Qy         316 ACATCATGCTCACCAGTCCTCGGACGCGTCCTGGACGCGTCCCCTCACTACCTCTTCG 375
               ||||||||||||||||| |||||||||||||| || |||||| ||||||||||||||
Db         765 ACATCATGCTCACCAACGTCCTGACGCGCCGCGGTCTCCTCCTACTACCTACCTATTCG 824

Qy         376 GCTTCGCCTTCGCTTTCGGCTCCCCCTCCCAAGGCTTCATCGTGAAACATTTCTTCGGCC 435
               ||||||||| |||| |||||||| ||||||||||||||||||||||||||||||||||||
Db         825 GCTTTGCATTCGCTTTCGGCTCGCGGCCCCTCATCGGCCGCCACTTCTTCGGCC 884

Qy         436 CTTCATCCTCCTACGACTACAGCTACTTCCTCCTACCAAT 486
               |||||||
Db         885 TACCAGATTACCCAATGGGCTCCTCCTCCCCTGGGGACTACACAGCTCTTCCTCTTCCTACCAGT 944

Qy         487 GGGGCCTTCGCCATCGCCGCGGCCATCACCACCAGC-GGAAGCATCGCGGACGCACACAG 545
               |||||||||||||||||||||||||||||| |||||| ||||||||||||||||||||||
Db         945 GGGGCCTTCGCCATCGCCGCGGCCATCACCACCAGCCTCCATCGCCGAGAGAACACAG 1004

Qy         546 TTCGTGGCCTATCTCATCTCACCGGCTTGTCTATCCGGTGGTCTCC 605
               ||||||||||||||||| ||| |||||||||||||
Db         1005 TTCGTGGCTTACTTATCTCATCTCTTTCTTCTTAACCGGTTCGTTCGTTCGTTACCCATCGTTTCG 1064

Qy         606 CACTGGTTCTGGTCCCAGACGCGGCTGGGCTCTGCTTTAAGATCACCG---ACCGGCTA 662
               |||||||||||||||||||||||| |||||||||| |||||||||| ||||||
Db         1065 CATTGGTTCTGGTCCTCAGACGGTTGGGCAGCGCGACTCGTAGCCACGGAAATGTTTTA 1124

Qy         663 TTTTCCACGGCGTAATAGACTTCGCCGGTTCCGGCGGTCCGAGTAGTCCTCGGCGGAATA 722
               ||||||||||||||||| |||||||| |||||| |||||| ||||| |||||||||||||
Db         1125 TTTCGGGTCTGGAGTCATCGACTTCGCGGGGTCAGGCGTTGTTCACAGGCGTTGGCGGGATA 1184

Qy         723 GCCGGCCATGGGGAGCGGCTGATCGAGAATGGACGTTTCGATCATGCAGGA 782
               ||||| |||| |||||||| || || |||||| | ||||||||||||||||
Db         1185 GCGGTCCTGTGTGGGGGGGCTTTAACGGGGGCTTTAGTTGGCTTGTCGACCGGTTCCGGC 1244

Qy         783 CGAGCTGGCCTTGCGAGGCCACAGCCGCGTCCTTAGTAGTCCTCGGGAACCTTCTTGCTT 842
               || |||||||||||||| |||| |||||||| |||||||||| ||||||||||||||||||
Db         1245 CGGTCGGTTGCTTACGTGGCCACAGCCGGCGTCTTTAGTTGTCGTTGTTTTGTTA 1304

Qy         843 TGGTTCGGTTGGTACGGATTTAACCCCGGTTCATTAACAAAATCCTACTTACTTACGGT 902
               ||||||||||| || |||||||||| ||| |||| |||| || |||
Db         1305 TGGTTCGGCTGGTACGGCTTCAACCCGGTTCGTTTCGTTTTCTGACAATAGCCAAGGGGTATGGA 1364

Db         544 GGTCAGTTGATGGATGGGCCAGCCCGTTCCGTACCGATGGAGATTTGCTTTTCAGCACCG 603

Qy         673 GGGTAATAGACTTCGCCGGTTCCGGCGTAGTCCACATGGTCGGCCGGAATAGCCGGCCTAT 732
               |||| ||| ||||  |||   |||||||| ||||| || |||  |||||||||| ||||||
Db         604 GAGCGATAGATTTCGCTGGTTGTTCATATGGTCGGACGTGTTGTTCATGGTATCGCTGGACTCT 663

Qy         733 GGGGAGCGGCTGATCGAAGGCCCAAGAAATGGGACGTTTCGATCATCGAGGACGAGCTGTGG 792
               ||||| |||||||||||| |||  ||| |||| ||||| |||||||| ||||| |||||||
Db         664 GGGGTGCGCTCATCGAAGGTCCACGACTTGGCCCGGTTCGGATAAACGGAGGGCCGGTGCCATCG 723

Qy         793 CCTTGCGAGGCCACAGCGCGGTCCTTAGTAGTCCTTGGGAACCTTCTTGCTTTGGTTCGGTT 852
               ||||||||||||||| |||   ||||||| ||||||| ||||||| |||| |||||| |||
Db         724 CTCTTCGTGGCCACTCGGCGGTCACTTGTTGTCCTTGGAACATTCCCTCCTCTCGGTTTGGAT 783

Qy         853 GGTACGGATTTAACCCCGGTTCATTCATTTAACAAAATCCTACTTACTTACGGTTAACTCAGGAA 912
               ||||||||||||||| ||||||    ||   ||||| |||||||| ||||||||||  || ||
Db         784 GGTACGGATTTAACCCCGGTTCCTTCAACAAGATCCTAGTCACGTTACGTGACGGACGA---CAGGCA 840

Qy         913 ATTACTACGGTCAATGAGACGCGGTTGGCAGAACCGCGGTCACCACTCCTAGCCGGGGT 972
               ||||| |||| ||||||||||||||| |||| |||||| ||| | |||| ||| |||||
Db         841 CATACAACGGCCAGTGAGGTGGAGCGCGGTCGGACGGACAGCAGTCGTCACAACAAGCTTAGCTGGCT 900

Qy         973 CAACAGCTGCCTTGACACGCTATTCGGTAAACGGGTGATATCCGGTCACTGGACGTGA 1032
               ||||||||| || ||||||   ||||     |||    |||  ||  |||||||||
Db         901 GCACCGCGCGCTGACAACCTATTTGGGAAAACGTCTACTCTCCGGACACATTGGAACGTCA 960

Qy         1033 CCGATGTCTGCAACGGGCTGTGTAGGCGGTTCGCGGCGGTTCGGCCCGATAAACAGCCGGTTGCTCCCGTGG 1092
                || ||   |||| |||||||| |||| ||| || ||| |||  ||||||||| ||| ||| || |||||||
Db         961 CTGATGTATGCAACGGGCCTCCTCGGAGGGTTTGCAGCCATAACTGTGTGGCTGCTCGTCG 1020

Qy         1093 TTGAGCCATGGGCAGCCATCGTATGCGGTTTGTTGCTTCTATAGTATTAATAGCTTGCA 1152
                ||||||||||||||||||| || |||||| |||| ||||||| |||| || ||||||
Db         1021 TTGAGCCATGGGCAGCCATCGTGCGGCGTTCGGTGGCCCCTAGTCCTCCGGATGCA 1080

Qy         1153 ACAAATTAGCAGAGAAGGTTAAGTTGACGATCCTCTGGAGGCGGCGCAGTTGCACGGTG 1212
                |||||| |||||| |  || || ||| |||| |||||||||||| ||| ||||| |||
Db         1081 ACAAGCTCGCTGAGAAGCTCAAATACGACACCCTCTTGAGCCAGCAGCACAACTACACGCGGTG 1140

Qy         1213 GGTGTGGCACGTGGGGGGGTGATATTCACGGGGTTGTTCCGAAAAAGGAGTATGTGAAGG 1272
                |||||||||| || |||| |||||||| |||  ||| || |||  ||| || || ||
Db         1141 GTTGCGGTGCGTGGGGACTAAATATTCACGGCTCTTGCTCGCAAGAAAAGTACTTGAACC 1200

Qy         1273 AGGTTTACGGGTTG-----GGGAGGGGGCCACGGGTTGCTCATGGGGGGTGGTGGAAGT 1326
                ||||||| ||||||      ||| |||||||||| || |||||||| || |||||||
Db         1201 AGATTTACGGCAACAAACCCGGAAGGGCCACACCGGTTTGTTATGGGCGGTTGGAGGAAAAAC 1260

Qy         1327 TGCTGGCGGCCGCACGTGATTCAGATTCTGGTGATTGCTGGGTGGTTAGTGCGACCATGG 1386
                ||||||| ||||| ||| ||||||||||| |||| |  ||||||||||| ||  ||||||
Db         1261 TACTTGAGCTCAGCTGATTCAGATCATTGTGATCACGGGTTGGGTAAGTGCGACCATGG 1320

Qy         1387 GACCCTTGTTTTGGGGGTTGAATAAACTGAAGTGTTGAGGATTTCTTCAGAGGATGAGC 1446
                ||||| ||||  ||| | ||||  ||||| ||| | |||||||| |||||||||||||
Db         1321 GGACACTTTTTCCTCATCTCAAGAAAATGAAATTGTTGCGGATATCGTCCGAGGATGAGA 1380

Qy         1447 TTGCGGGGATGGACATGACTCGCCATGGAGCGTTTGCTTATGCTTATGGATGGATGAGA 1506
                ||||| ||||||| |||||||||| |||||||| || || ||| ||||||||||| ||
Db         1381 TGGCCCGGTATGGATGGATGACCAGGCACGGTGGTTTTGCTTATATGACTTTGATGATGATG 1440

Qy         1507 CGCACAAGCACATGGGATGCAGTTGAGGAGGGTTGGGCCCAAGCGGTCTTCCACACCCACCA 1566
                ||||| ||||| || |||||||| |||| || |||||||| |||| | |||||||||||||
Db         1441 AGTCTCACAAGCACCATTCAGCTTAGGAGAGTTGAGCCACGACGATCCTCCTCTCCTCCTCTCGGTG 1500

Qy         1567 CTGAT 1571
                |||||
Db         1501 CTAAT 1505

RESULT 7
US-10-424-599-54229
; Sequence 54229, Application US/10424599
; Publication No. US2004031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
```

```
Best Local Similarity  70.3%;   Pred. No. 9.3e-177;
Matches  979;  Conservative   0;  Mismatches  402;  Indels  12;  Gaps  2;

Qy     124 CAAACACCACGAGACGCTTCCGCCGCCCGCCTCCCTATCTGCGGGCCATTCGCCGCCGTGG 183
               ||||||  |||||||  |||||| ||  |||| ||| |||||||  |||| | ||||||
Db      62 CTAACTCAACATCTTCCCTCGCCGCGGCAACCTTTTTATGTTCCCAAATTTCAAACATCT 121

Qy     184 ACAGCAAGTTCGTCGACACGGCCTTCGCCGTCGACAACACCTACCTCCTCTTTTCCGCCT 243
               |||| ||| || ||||||||| ||||| || |||  ||||  |||||||||||||||
Db     122 CCAACAAACTCTCGACACAACTTATGCCGTCGACAACAACAGTATCTCCTCTTCTCCGGCT 181

Qy     244 ACCTCGTTTTTCTATGCAGCTCGGCTTCGCCATGCTCGTCGCCGCCGCTCCGTCCGGGCCA 303
               |||| ||| | ||| ||| |||||||| ||  |||||| || |||||| |||||||||
Db     182 ACCTTGTCTTTGCCATGCAGCTCGGTTTCGCTATGCTTTGTGCTGATCAGTCGAGCCA 241

Qy     304 AGAACACCATGAACATCATGCTCACCAAGTCTCGGACGTCTGGACGCGTCGCCGGCGCAG 363
               |||| |||||| |||||||| |||||| |||| |||| ||| ||||||||||||| ||
Db     242 AGAACACTATGAACATCATGCTTACCCAATGTCCTTGATGCTGCCCGCTGGAGCCCATCTCTT 301

Qy     364 ACTACCTCTTCGGCTTCGCTTCCGCTCCCCCTCCCAACGGCTTCATCGGTAAAC 423
               |||||||||||| ||| | |||| | ||||||||    ||||||| ||||||
Db     302 ACTACCTCTTCGGATTCGCATTCGCCTTTGGTACACCTTCCAACGATTCATCGGTCGCC 361

Qy     424 ATTTCTTCGGCCTCAAGGACATCCCTT--------CATCCTCCTACGACTACAGCTACT 474
               | ||| || ||| || |||| ||||             |||| | ||| |||||||
Db     362 ACCATAGCTTCTTCGCTTTAAGCTCTTACCCTGAACGCCCGCTTCGACTTCAGCTTTT 421

Qy     475 TCCTCTACCAATGGGCCTTCGCCATCGCCGCCGGCATCAACCAGCGGAGCATCGCCG 534
               |||||||| ||||||  |||||||| ||||| ||||| || |||| ||||||||||
Db     422 TCCTCTACCAATGGGCTTTTGCCATAGCGCGCCGGAATCACTAGCGGTTCCATCGCCG 481

Qy     535 AACGCACACAGTTCGTCGCCTATCTCATCTCATTCCTCCTTCCACCGGCTTCGTCTCATC 594
               |||| ||| ||| |||||| || || |||| | | |||| || |||| ||| | ||||
Db     482 AGCGAACGCAATTCGTTCGTTCTTGCTTATCTATCTTTCTTGACCGGTTTTGTTTACC 541

Qy     595 CGGTGGCTCCCACTGGTTCTCGTCCCCAGACGGCTGGGC---CTCTGCCTTAAGATCA 651
               |||||| || |||||||| ||||| ||| | ||||| ||    | |||||||||||
Db     542 CGACAGTTCGCACTGGTTCTCGGTCAAGTGATGGGCTAGCGGCGGTAGCGGTCTGACA 601

Qy     652 CCGACCGGCTATTTCCACCGGGTAATAGACTCGCCGGTTCCGGCGAGCCCAAGAATGGAC 711
               ||||  || |||||| || ||| |  || |||| ||||||| ||||||| ||||||
Db     602 ACAATCTCTTGTTTGGCTCAGGTGCTATTGATTTCGCAGGTTCAGGTTCACATGG 661

Qy     712 TCGGCGAATAGACGGCCTATGGGAGCGCTGATCGAAGGCCAAGAATGGACGTTTCG 771
               ||| |||| || || |||||| |||| ||| || |||| ||| || |||||||||||
Db     662 TAGGTGGAATTGCCGGTTTTATGTGGAGCGTTAGTCGAGGCGTTGAGGACCAAGAATAGGTAGATTTG 721

Qy     772 ATCATGCAGGACGAGCTGTGGCCTTGCGAGGGCCACAGCCGCGTCCTTAGTAGTCCTCGGAA 831
               || ||||| || || ||||| || |||||||| || ||| ||| |||||| |||||||||
Db     722 ACCGGTCAGGCCCGGTCCGGGCTTTACGTGGTCACAGTGCATCCCCTTGCTCGTCGTCGGTA 781

Qy     832 CCTTCTTGCTTTGGTTCGGTTCGGTTACGGATTTAACCCCGGTTCATTTAACAAAATCCTAC 891
               |||||||||| ||| || |||| |||| || |||||||| | ||||| | ||||| ||||
Db     782 CCTTCTGTGTGTTTGGATGGTATGGGTTTAACCCCGGTTCCTTTTTAACCATTCTTA 841

Qy     892 TTACTTACGGTAACTCAGGAAATTACTACGGTCAATGAGCTCGGTTGGCAGAACCGCGG 951
               ||||      |  ||||||||||| |||| |||   ||| |||| ||||||| ||||||
Db     842 AAGGCTACGACAAGTCTCGGCCATTATATGGTCAATGAATGGCGTGTAGGTCGCACCGCGG 901

Qy     952 TCACCACTACCCTAGCGGGGTCAACAGCTGCCTTGACCACGCTATTCGGTAAACGGGTGA 1011
               |||||||| |||   ||||  |  |     |||| |||||| ||| ||||||||| |||
Db     902 TCACCACAACGCTTTCTGGCTGCGACCGCGTTGACTACTCTATTCAGTAAAACGGCTTT 961

Qy    1012 TATCCGGTCACTGGAACCTGACCGATGTCTGACAACGGGCTGTTAGGCGGTTTCGCCGCGA 1071
               ||||||||| | || ||| ||  ||||||| || | ||| ||  ||| |||||||| |||
Db     962 TAGCAGGTCATTGGAACGTTATTGACGTATGCAACGGACTTCTAGGCGGGGTTTTGCAGCTA 1021

Qy    1072 TAACAGCCGGTTGCTCCGGGTTGAGCCATGGCAGCCATCGTATGGCCGGTTTTTGTTGCTT 1131
               |||| |||| |||||| | |||| |||||||| || |||| ||||  || ||||| |||
Db    1022 TAACCTCCGGATGCCCGTGGTGGAGCCCGGTGCTATAGTATGTGGGTTTTGGCAT 1081

Qy    1132 CTATAGTATTAATAGCTTGCAACAAATTAGCACAGAGAAGAGGTTAAGTTCGACGATCCTCTGG 1191
               |  |||||| |  |||| ||| ||     ||||    |   |||| |    |||| | |||
```

```
Qy     903 AACTCAGGAAATTACTACGGTCAATGGAGCGCGGTTGGCAGAACCGCGGTCACCACTACC 962
               |||||||||| |||| ||||||||  ||||||||   |||||||||||||| |||||||
Db    1365 AGTGGAGGTATTAT--GGTCAATGAGCGGCTATAGGGAGGACAGCTGTCACGACGACA 1421

Qy     963 CTAGCGGGGTCAACAGTTGCCTTGACCACGCTATTCGGTAAACGGGTGATATCCGGTCAC 1022
               ||| |||| || || ||| |||| |||||||| |||||| ||||||||| || ||||||
Db    1422 TTGGCTGGGAGCACTGCGGCGTCTCGACGACGTTGTTCAGCAAGCCGGTTATTGGCTGGCCAC 1481

Qy    1023 TGGAACGTGACCGATGTCTGCAACGGGCTGTTAGGCGGTTCGCGGCGATAAACAGCCGGT 1082
               |||||||||||| || || ||||||||| ||||  ||| |||||| ||| ||||||||
Db    1482 TGGAACGTGATTGACGTGTGTAACGGCCTCGGCCTTGGCGGGGTTTCGCTGCCATTACATCGGGC 1541

Qy    1083 TGCTCCGTGGTTGAGCCATGGCAGCCATCGTATGCGGTTTTGCTTCTATAGTATTA 1142
               ||| |||||| |||| ||||| || ||| | | ||| ||||| || || || |||
Db    1542 TGTGCCGTTGTGGAACCGTGGCCGCGATTGTGTGTGGGTTTGCTGCTATGCTTG 1601

Qy    1143 ATAGCTTGCAACAAATTAGCAGAGAGGTTAAGTTCGACGATCCTCTGGAGGCGGCGCAG 1202
               || |||| || || || |||| |||||  ||| ||||| ||| ||  |||||| ||||
Db    1602 ATTGGGCTAATAAGCTTGCCGCGAAGGTAGAGTACGATGATCCGTTGGAGGCGCGCAG 1661

Qy    1203 TTGCACGGTGGTGTGGCACGTGGGGGGTGATATTCACGGCGTTGTTCGCAAAAAAGGAG 1262
               |||||||| |||| |||| ||| ||||||||||||| | ||  || |||||||| ||||
Db    1662 CTTCACGCGCGGGGTGCGGAGGCGTGGGGGGGTTTCTTCACGGGATTGTTTGCGGAAGAAAGTG 1721

Qy    1263 TATGTGAAGGAGGTTTACGGGTTTGGGGAGGGGCACGACGGTTGCTCATGGGGGTGGTGGG 1322
               | |||| |||||||||||| |||||||| || ||    ||||| || ||||||||||| |||
Db    1722 TACGTGGAGGAGATTTACGGTGTTTGGAAGGCCCGTTCGGGGCTTTGAATGGGTGGCGGAGGG 1781

Qy    1323 AAGTTGCTGGCGCGCACGTGATTCAGATTCTGGTGATTGCTCGGGTGGGGTTAGTGCGACC 1382
               |||||| || ||||||| ||| |||| || ||||||  | | ||||| ||| ||| ||||
Db    1782 AGGCTGCTGCGCGCGCAGGTGATTCAGATTCAGGTGGTGCGCGGGTGGGTGGGTTACGGCGACC 1841

Qy    1383 ATGGGACCCCTTGTTTTGGGGGTGAATAAACTGGAGCTGTTGAGGATTTCTTCAGAGGAT 1442
               |||| | ||| |||||| || |||  ||| || || ||| |||| ||| || ||||||||
Db    1842 ATGGGCGCGTTGTCGTTGTTGCTTCATGAGATGAAACTGTTGAGAATTTCGAGGAGGGATGAT 1901

Qy    1443 GAGCTTGCGGGGATGGACATGCTCGCCACTGGACATGGAAGCTTTGCTTATGCTTATGAGGATGAT 1502
               ||||||||| |||| |||||||  | | |||| | ||| |||| |||| ||||| ||| |||
Db    1902 GAGACTGCGGGGATGGATTTGACGAGGACATGTGGGGTTCTTATGCTTATGCATACCATGATGAT 1961

Qy    1503 GA 1504
               ||
Db    1962 GA 1963
```

Query Match   41.4%; Score 705.8;   DB 9; Length 1545;

RESULT 8
US-09-938-842A-1490

; Sequence 1490, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIP1300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1490
; LENGTH: 1545
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-1490

```
Db   1082 CATGGGTTTTAATCGGATTAACTTGCTTGCCAAGAAACTTAAATATGATGACCCACTCG 1141

Qy   1192 AGGCGGCGCAGTTGCACGGTGTGGCACGTGGGGGGGGTGATATTCACGGCGTTGTTCG 1251
          ||||||  |||||| |||   ||| |||| ||| ||||    ||||| || |||||||
Db   1142 AGGCTGCTCAGCTCCACGCGTGATGTGGAGCATGGGAGATTAATCTTTACCGGGCTGTTCG 1201

Qy   1252 CAAAAAGGAGTATGTGAAGGAGGTTTACGGGTTGGGAGGGCCACGCGGGTTGTCATGG 1311
          ||||| |||||| ||  |||  ||| ||||| ||| |||   ||||| |||||||||||
Db   1202 CAAGGAAAGAATACGTTAACGAGATTCTCCGGTGATAGGCCTTCATGG 1261

Qy   1312 GGGGTTGGTGGGAAGTTGCTGCGCGGCACGCAGTGATTCAGATTCGGTGATTGCTGGGTGG 1371
          ||||   ||||| ||| ||||  ||||| || |||| ||||| ||||| ||||||||||
Db   1262 GCGGGGGAGGGAAAACTGCTCGCGCCGCAGATCGTTCAGATTATTGTGATCGTTGGGTGGG 1321

Qy   1372 TTAGTGCGACCATGGGACCCTTGTTTGGGGTTGAATAAACTGAAGCTGTTGAGGATT 1431
          ||||||||| ||| |||||| |||| ||||||||||| ||  |||||||||||||||
Db   1322 TGACGGTAACTATGGGACCGTTGTTTTATGGGTTACATAAGATGAATCTTTGAGGATAT 1381

Qy   1432 CTTCAGGAGGATGAGCTTGCGGGGATGGCATGATGACTCGCATGAGGCGTTGCTTATGCTT 1491
          ||||| |||||| |||| |||| |||   ||||| |||||||   ||||  |||||||
Db   1382 CAGCAGAAGAAGATGAGATGGCAGGAATGGGCATGATGACACGTCATGGAGGATTGCTTACGCAT 1441

Qy   1492 ATGAGGATGATGA 1504
          ||   |||||
Db   1442 ACAATGACGAAGA 1454
```

RESULT 9

US-09-938-842A-1490

US-09-938-842A-1490

Query Match 41.4%; Score 705.8; DB 11; Length 1545;
Best Local Similarity 70.3%; Pred. No. 9.3e-177;
Matches 979; Conservative 0; Mismatches 402; Indels 12; Gaps 2;

```
Qy    124 CAAACACCACAGACGCCTCCGCGCGCCTCCCTATCTGCGGCCATTTCGCGCCGCCGTGG 183
          ||||| ||   |||||||  || ||| ||||  ||  |||  ||||  ||||| |  |
Db     62 CTAACTCAACATCTTCCCTGCCGCCGCAACCTTTTATGTTCCCAAATTCAAACATCT 121

Qy    184 ACAGCAAGTTCGTCGACACCGGCCTTGCCGTCGACACACACTTACCTCCCTCTTTTCCGCCT 243
          ||||||||| ||| ||||  |||| |||||||| |||  || || || |||||||||||
Db    122 CCAACAAACTCTCCGACACAACTTATGCCGTCGACAACACGTATCTCCTCTTCTCCGGCCT 181

Qy    244 ACCTCGTTTTTCTATGCAGCTCATGTCTCCGGACGCTGCCGCCGCCGCCGCCGTCCGCGCCA 303
          ||||||||||| ||||||| ||  || |||| ||| || ||| |||||| |||||||||||
Db    182 ACCTTGTCTTTGCCATGCAGCTCGGTTCGCATGCTTGTGCTGCGCGCCGCCGCGAGCCA 241

Qy    304 AGAACACCATGAACATCATGTCACCAACGTCTCGGACGCTGCCGCCGCCGGCTCTTCT 363
          ||||||| |||||| |||| | | ||| ||| ||| ||||| |||||| |||||||
Db    242 AGAACACTATGAACATCATGTCCTTACCAATGTCCTTACCAATGTCCTTGATGCTCCGCGTCGGAGCCATCTCTT 301
```

Qy 1432 CTTCAGAGGATGAGCTTGCGGGATGACATGACTCGCCATGAGGCTTGCTTATGCTT 1491
         | ||| |||||||| ||| ||||||||||| || |||||||| ||||||||| |
Db 1382 CAGCAGAAGATGAGATGGCAGGAATGACATGACACGTCATGGAGGATTTGCTTACGCAT 1441

Qy 1492 ATGAGGATGATGA 1504
         || ||||||||||
Db 1442 ACAATGACGAAGA 1454

RESULT 10
JS-09-938-842A-1887
; Sequence 1887, Application US/09938842A
; Patent No. US2002160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIP1300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1887
; LENGTH: 1497
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
JS-09-938-842A-1887

Query Match   41.1%; Score 700.4; DB 9; Length 1497;
Best Local Similarity 69.7%; Pred. No. 2.5e-175;
Matches 995; Conservative 0; Mismatches 421; Indels 12; Gaps 3;

Qy 107 GGGCCAACTTCTCGGCCCAAACACCAGACGCCTCCGCCGCCTCCCTTATCTGCGG 166
         ||||||||||||||||||| || |||||||| ||||| | |||||||||||||
Db   27 GGCCGATCTCGCCACCCTACTTGGCCCCAAGCGCCACGGCGCCGACTACATTCGCGG 86

Qy 167 CCATTTCGCCGCCGTGGACAGCAAGTTCGTCGACACGGCCTTCGCCGTCGACAACACCTA 226
         ||||||||||||||| ||||||| ||||| |||||||| |||||||||||||
Db   87 CCAATTAGGCACCGTTAACACAAGTTCACCGATGCAGCCTTCACCGTTCGCCATAGACAACACCTA 146

Qy 227 CCTCCTCTTTCCGCCTACCCGTTTTTCATGCAGCTCGCTTGCTGCCATGCTCTGCGC 286
         |||||||||||| || |||| ||||| || |||||| |||||||||||||
Db 147 CCTCCTCTTCTCTGCCTACCTTGTCTTCGCCCATGCAGCTCGGCTTCGCCATGCTTTGTGC 206

Qy 287 CGGCTCCGTCCGGCCAAGAACACCATGAACATCATGCTCACCAACGTCCTGGACGCTGC 346
         ||||| ||||||||||| ||||| | | |||||| || |||||||||||||
Db 207 TGGTTCTGTTAGAGCCAAGAATACGATGAACATCATGCTTACCAATGTCCTTGACGCTGC 266

Qy 347 CGCCGGCGGCCTCTTCTACTACCTTCTCGGCCTTCGCCTTTCGCCTCCCCCTCCAA 406
         ||||| ||| ||||||||| || |||||||| |||||| |||||
Db 267 AGCCCGAGGACTCTTCTACTATCTCTTTGGTTACGCCTTTTGCCTTTGCCTTTGGAGGATCCTCCGA 326

Qy 407 CGGCTTCATCGGTAAACATTCTTCGGCCTCAAGGACATCCTTCATCCTTCCTACGACTA 466
         ||||||||| || ||| | || |||| || |||||| || || ||||||||||
Db 327 AGGGTTCATTGGAGACACAACTTTGCTCTTAGAGACTTTCCGGACTTCCACAGCTGATTA 386

Qy 467 CAGCTACTTCCTCTACCAATGGGCCTTGCCATCGCCGCCGGCAATCACCAGCGGAAG 526
         ||||||| ||| |||||||| ||| |||||||||||||||| ||||| |
Db 387 CTCTTTCTTCCTCTACCAATGGCGTTCGCAATCGCGGCCGCTGGAATCACAAGTGGTTC 446

Qy 527 CATCGCCGAACGACACACAGTTCGTGGCCTATCTCATCCTCCTTCCTCCTCACCGGCTT 586
         |||||||| || || |||||||| |||| |||| | ||||||| |||||||
Db 447 GATCGCAGAGAGGACTCAGTTCGTGGCTTACTTGATATACTCTTCTTTCTTTAACCGGATT 506

Qy 587 CGTCTATCCGGTGGTCGTCTCCCCAGACGGCTGGGCCTCTGCCTTTAA 646
         || ||| | |||| |||||||||||| || ||||| || ||||| ||
Db 507 TGTTTACCCGGTTGTCTCTCACTGGTTTCCACGGCTATTTCCACGGCTTCCTTCG 566

Qy 647 GATCAC---CGACCGGCTATTTTCCACCGGCGTAATAGACTTCGGCCGGTTCCGGCGTAGT 703
         ||||| ||| | |||| || | ||||||| || |||| || ||||||||||
Db 567 TTCAGCGGATGATCGTTTGTTTAGCACCGGACCATTGACTTTGCTGGCTTCCGGTGTTGT 626

Qy 704 CCACATGGTCGGCGGAATAGCCGGCTCATGGGAGCGGCGTGATCGAAGGCCCAAGAATGGG 763
         |||||||||| ||| ||| | |||||||||| || ||||||||||
Db 627 TCACATGGTTGGTGGCATAGCAGTTTATGGGGTGCTCTTCTTATTGAAGGTCCGTCGTGG 686

Qy 764 ACGTTCGATCATGCAGGACGAGCTGTGGCCTTGCGAGCCACAGCGCGTCCTTAGTAGT 823
         ||||||||||| ||||||||| || | || ||||||||||| || | |||||||
Db 687 TCGGTTCGAGAAGGTGTCGCGCTATTGCTCGCGGCCGCCACTCGCCTAGT 746

Qy 824 CCTGGGAACCTTCTTGCTTTGGTTGGTACGGATTAACCCGTCATTTAACAA 883
         |||||||||||||||||| |||||||||| || || |||||||||||
Db 747 CTTAGGAACCTTCCTCCTATGGTTTGGATGGTATGGTTCAACCCCGGTTCCTTCACTAA 806

Qy 884 AATCCTACTTACTTACGGTAACTCAGGAAATTACTACGGTCAATGGAGCGCGGTTGGCAG 943
         ||| || || | || |||||||||| |||||||||
Db 807 GATACTCGTTCCGTA---TAATTCTCGGTTCCAACTACACGGCCAATCGGCGGAATCGGCCG 863

Qy 944 AACCGCGGTCACCACTACCCTAGCGCGGGTCAACAGCTGCCTTGACCACCGTATTCGGTAA 1003
         |||| ||||||| || || |||||||||||||||
Db 864 TACAGCGGGTTAACACCACACCTCAGCAGCTCTAACCACACTCTTTGGTAA 923

Qy 1004 ACGGGTGATATCCGGTCACTGGACAAGTGGACCGATGTCTGCAACGGCTGTGTTAGGCGGTTT 1063
         | ||| |||||||||||||||||||||||
Db 924 ACGTCTTCCTATCAGGCCACTGAAACTGTAACGACGTTTGCAACGGGTTACTCGGTGGGGTT 983

Qy 1064 CGCGGGCGATAACAGCCCGGTTGCTCCGTGGTTGAGCCATGGGCAGCCATCGTATGCGGGTTT 1123
         ||||| | ||||||||||||||||||||
Db 984 TGCGGCCCATAACCGCAGGTTGCTCCGTCCGTGTAGAGCCGAGGCCATGTGTGTGCGGCTT 1043

Qy 1124 TGTTGCTTCTATAGTATTAATAGCTTGCAACAAATTAGCAGAGAAGGTTAAGTTCGACGA 1183
         |||||||||||||||||||||||||||
Db 1044 CATGGCTTCTGTCGTCGTCCTTATCGGATGCAACAGCTCGCGGAGCTTGTACAATATGATGA 1103

Qy 1184 TCCTCTGGAGGCGGCGGCCAGTTGCACGTGGTGTGCACGTGGGGGTGTGATAATTCACGGC 1243
         ||||| ||||||||||||||||||||
Db 1104 TCCACTCGAGGCAGCCCAACTACATACGAGGGGTGTGGCCGTGGGGGTTGATATTCGTAGG 1163

Qy 1244 GTTGTTCGCAAAAAGGAGTATGTGAAGGAGGTTACGG------GTTGGGAGGGCGCA 1297
         ||||| || ||||||||||||||||||||
Db 1164 ATTGTTTGCCAAGAGAAGTATCTAAACGAGGTTTATCGGCGCCACCCCGGGAAGGCCATA 1223

Qy 1298 CGGGTTGCTCATGGGGCGTGGTGGGAAGTTGCTGGCGGCGCACGTGATTCAGATTCTGGT 1357
         || || || |||||||||||||||||||
Db 1224 TGGACTATTTATGGGCGGAGGAGGAAGCTGTTGGGAGCACAATTGGTTCAAATACTTGT 1283

Qy 1358 GATTGCTGGGTGGGTTAGTGCCATGGGACCCTTGTTTGGGGGTTGAATAAAACTGAA 1417
         ||||| |||||||||||||||||||
Db 1284 GATTGTAGGATGGGTTAGTGCCACAATGGGACACTCTTCTTCATCCTCAAAAGGCTCAA 1343

Qy 1418 GCTGTTGAGGATTTCTTCAGAGGATGAGCTTGCGGGGATGGACATGACTCGCCATGGAGAGG 1477
         ||||||||||||||||||||||||
Db 1344 TCTGCTTAGGATCTCGGACGCAGCAGCATGAAATGCAAGGATGGATAGTGACACGTCACGGTGG 1403

Qy 1478 CTTTGCTTATGCTTATGAGGATGATGAGAGACCGCACAAGCCATGGGGATGCA 1525
         |||||||||||||||||||||||
Db 1404 CTTTGGTTATATCTACCATGAAATGATGATGAGTGTCTCATAGAGTGGA 1451

RESULT 11
US-09-938-842A-1887
; Sequence 1887, Application US/09938842A
; Publication No. US20040009476A9
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong

```
;   TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
;   TITLE OF INVENTION: SAME, AND METHODS OF USE
;   FILE REFERENCE: SCRIP1300-3
;   CURRENT APPLICATION NUMBER: US/09/938,842A
;   CURRENT FILING DATE: 2001-08-24
;   PRIOR APPLICATION NUMBER: US 60/227,866
;   PRIOR FILING DATE: 2000-08-24
;   PRIOR APPLICATION NUMBER: US 60/264,647
;   PRIOR FILING DATE: 2001-01-16
;   PRIOR APPLICATION NUMBER: US 60/300,111
;   PRIOR FILING DATE: 2001-06-22
;   NUMBER OF SEQ ID NOS: 5379
;   SEQ ID NO 1887
;   LENGTH: 1497
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
US-09-938-842A-1887
```

```
Query Match    41.1%;    Score 700.4;  DB 11;  Length 1497;
Best Local Similarity   69.7%;   Pred. No. 2.5e-175;
Matches 995;  Conservative  0;  Mismatches 421;  Indels 12;  Gaps 3;
```

```
Qy    107  GGCCCAACTCTTCGGCCCAAACACCAGAGCGCCTCCGCGCCGCCTCCCTTATCTGCGG  166
                ||||||  ||| ||||||  ||||||   ||||||||||||||||||||
Db     27  GGCCGATCTCGGCCACCCTACTTGGCCCCAAGCCACGGCGGCCGACTACATTTGCGG   86

Qy    167  CCAATTCGCGCCGGTGGACAGCAAGTTCGTCGTCGACACGGCCTTCGCCGTCGACAACACCTA  226
                ||||||| ||||| ||||||||  |||||||  |||||||||||||||||||
Db     87  CCAATTAGGCACCGTTAACAACAAGTTCACCGATGCAGCCTTCGCCATAGACAACACCTA  146

Qy    227  CCTCCTCTTTTCGCCCTACCTCGTTTTTCTATGCAGCTCGGCTTCGCCATGCCTCGCGC  286
                ||||||||| |||||| |  ||||||||| |||||||||||| ||||
Db    147  CCTCCTCTTCTCTGCCTACCTCGTTTGTCTTCGCCATGCCTATGCTTGTGTGC  206

Qy    287  CGGCTCCGTCCGCGCCAAGAACACCATGAACATCATGGTCTCCTGGACGCTGC  346
                ||||||||  |||||||||||||||||||||||||||| ||||||
Db    207  TGGTTCTGTTAGAGCCAAGATACGATGAAACATCATGCCTTACCAATGTCCTTGACGCTGC  266

Qy    347  CGCCGGCGGCCTCTTCTTACTACCTCTTCGGCTTCGGCTTTCGCTCCCCTCCAA  406
                ||||||||||||||||||   ||||||| ||||||||||||||||
Db    267  AGCCGGAGGACTCTTCTTACTATCTCTTTGGTTACGCCTTTGCCTTTGGAGGATCTCCGA  326

Qy    407  CGGCTTCATGGTAAACATTTCTTCGGCCTCAAGGACATCCCTTCATCCTCCTACGACTA  466
                |||||| |  ||| ||||||||  |||||  ||||
Db    327  AGGGGTCATTGGAAGACACAACTTTGCTCTTAGAGACTTTCCGACTCCCACAGCTGATTA  386

Qy    467  CAGCTACTTCCTCTACCAATGGGCCTTCGCCATGCCGCGCGCATCACCAGCGGAAG  526
                |||||||  ||||| ||  |||  |||||||||||||||||
Db    387  CTCTTTCTTCCTCTACCAATGGGCGTTCGCAATCGCGCCCGCTGGAATCACAAGTGGTTC  446

Qy    527  CATCGCCGACGCACACAGTTCGTCGGCCTATCTCATCTCCTCCTCCTCACCGGCGTT  586
                ||||| |||| | ||||||||| |||||| |||||||||||
Db    447  GATCCGCAGAGAGGACTTCAGTTCGTCGGCTTACTTCGATATACTCTTCTTTCTTAACCGGATT  506

Qy    587  CGTCTCATCCGGTGGTCTCCCACTGGTTCTCGGTCCCCAGACGACGGGCTCGGAGCTCCTCGCCTTTAA  646
                ||||| |||||||||||||||||| ||||||||| |||||||||||
Db    507  TGTTTACCCGGTTGTCTCTCACTGGTTTTGGTCCCCGGATGGATGGGCCAGTCCCTTTCG  566

Qy    647  GATCAC---CGACCCGCTATTTTCCACGCGGCTAATAGACTTCGCCGGTTCCGCGTAGT  703
                ||||     |||||||| |||||||| || |||||| ||||||
Db    567  TTCAGCGGATGATCGTTTGTTGTTAGCACCGGAGCCATTGACTTTGCTGCTCCCGGTGTTGT  626

Qy    704  CCACATGGTCGGCGAATAGCCGGCCTATGGGGAGCGGCTGATCGAAGGCCCAAGAATGGG  763
                ||||||||||| |||||| ||||||||||||||| ||||||
Db    627  TCACATGGTTGGTGGCATAGCAGGTTTATGGGTGTCTTATTGGAAGTCCTCGTCGTGG  686

Qy    764  ACGTTTCGATCATGCAGCGAGCAGTGTGGCCTTGCGAGGCCACAGCGCGTCCTTAGTAGT  823
                |||||||||||||||||||| |||||  ||||||| ||||||| |||||||||
Db    687  TCGGTTCGAAAGTGGTCGCGCATTGCTCGCGCGCCCACTCTGCCTCGCTAGTAGT  746

Qy    824  CCTGGGAACCTTCTTGCTTGTTGGTTCGGTTGTACGGATTTAACCCCGGTTCATTTAACAA  883
                |||||||| |||||||| ||||| |||||| ||||||||||||||||||||
Db    747  CTTAGGAACCTTCCTCCTATGGTTTGGATGGTATGGTTTCAACCCCGGTTCCTTCACTAA  806
```

```
Qy    884  AATCCTACTTACTTACGTAACTCAGGAAAATTACTACGGTCAATGGAGCGCGGTTGGCAG  943
                || ||||  | |||||||||||||||||||||||||||||
Db    807  GATACTCGTTCCGTA---TAATTCTGGTTCCAACTACGGCCACACGTCCTTGACACAGCTGCCTTCGGTAA  863

Qy    944  AACCGCGGTCACCACTACCCTAGCGGGGTCAACACAGCTGCCTTGACCACGGCGCTATTCGGTAA  1003
                |||| ||| |||||| | ||||  |||||| |||||||||||||||
Db    864  TACAGCGGTTAACACCACCACTCTCAGGATGCACAGCAGCTCAACCACACTCTTGGTAA  923

Qy   1004  ACGGGTGATATCCGGTCACTGGAACGTGACCGATGTCGCAACGGCTGTTAGGCGCGTT  1063
                |||||||||||||| ||||||||||||||||||||||||||| |||||||||
Db    924  ACGTCTCCTATCAGGCCACTGGAACGTAACCGGACGTTTGCAACGGGTTACTCGGTGGGTT  983

Qy   1064  CGCGGCGATAACAGCCGGTGCTCCGTGGTTGAGCCATGGGCAGCCATCGTATGCGGTTT  1123
                ||||||||| ||| |||||||| |||||| |||||||| |||||||||||||
Db    984  TGGGGCCATAACCGCAGGTTGCTCCGTCCGTGTAGAGCAGCCATGGGCAGCCAGATTGTGCGGCTT  1043

Qy   1124  TGTTGCTTCTATAGTATTAATAGCTTGCAACAAAATTAGACAGAGAAGGTTAAGTTCGACGA  1183
                ||| ||||||  ||||||||| || ||| ||||||||||||||
Db   1044  CATGGCTTCGTGTCGTCCTTATCGGATGCAACAAGCTCGCGGAGCTTGTACAATATGATGA  1103

Qy   1184  TCCTCTGGAGGCGGCGGCCAGTTGCACGGTGGGTGTGGCACGTGGGGGTGATAATTCACCGC  1243
                ||| |||||| |||||  ||| ||||||| ||||||||  ||||||||||
Db   1104  TCCACTCGAGGCAGCCCCAACTACATGGACGGTGTGGCGGTTGATATTCGTAGG  1163

Qy   1244  GTTGTTCGCAAAAAAGGAGTATGTGAAGGAGGTTTACGG------GTTGGGGAGGGGCA  1297
                |||| |||| | ||||||| |||||||||||||||||  |||||||||||||
Db   1164  ATTGTTTGCCAAAGAGAAGAGTATCTAAACGAGGTTTATGGGGCCACCCCGGGAAGGCCATA  1223

Qy   1298  CGGGTTGCTCATGGGGGTGGTGTGGGAAGTTGCTGGCGGCGCACGTGATTCAGATTCTGGT  1357
                |||  |||||||||| | |||||| |||||||| |||||  |||||
Db   1224  TGGACTATTTATGGGCGGAGGAGGGAAGCGTGTTGGGAGCACAATTGGTTCAAATACTTGT  1283

Qy   1358  GATTGCTGGGGTGGGTTAGTGCGACCATGGGACCCTTGTTTTGGGGGTTGAATAAACTGAA  1417
                || ||||||||||||| |||||| ||| |||||||||||| |||||||
Db   1284  GATTGTAGGATGGGTTAGTGCCACAATGGGAACACTCTTCTTCATCCTCAAAAGGCTCAA  1343

Qy   1418  GCTGTTGAGGATTTCTTCAGAGATGAGCTTGCGGGATGACATGACTCGCCATGGAGG  1477
                ||||||  | ||||| ||||||||||||||||  ||||||||||
Db   1344  TCTGCTTAGGATCTCGGAGCAGCATGAAAATGCAAGGGATGATATGACACGTCACGGTGG  1403

Qy   1478  CTTTGCTTATGCTTATGAGGATGATGAGACGGCCACAAGCATGGGATGCA  1525
                |||||||| |||| |  |||||||||||||| |||||
Db   1404  CTTTGCTTATATCTACCATGATAATGATGAGTCTCATAGAGAGTGGA  1451
```

```
RESULT 12
US-10-424-599-70091
; Sequence 70091, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
;   APPLICANT: La Rosa Thomas J
;   APPLICANT: Kovalic David K
;   APPLICANT: Zhou Yihua
;   APPLICANT: Cao Yongwei
;   TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
;   TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
;   FILE REFERENCE: 38-21(53223)B
;   CURRENT APPLICATION NUMBER: US/10/424,599
;   CURRENT FILING DATE: 2003-04-28
;   NUMBER OF SEQ ID NOS: 285684
;   SEQ ID NO 70091
;   LENGTH: 842
;   TYPE: DNA
;   ORGANISM: Glycine max
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT3847_34303C.1
US-10-424-599-70091
```

```
Query Match    28.6%;    Score 488.2;  DB 12;  Length 842;
Best Local Similarity   86.6%;   Pred. No. 4.2e-119;
Matches 538;  Conservative  0;  Mismatches 83;  Indels 0;  Gaps 0;
```

```
Qy    583  GCTTCGTCTATCCGGTGGTCTCCCACTCGTTCTCGTCCCAGACGGCTGGGCCTCCTGCCT  642
                ||||||||| ||||||||||| ||||||||||| |||||| ||||||||||||    |
Db      1  GGTTCGTCTACCCGTGGTCTCGTCCCACTCGTTCGTCGTCCGCGACGGCTGGGCCGGCCCA    60

Qy    643  TTAAGATCACCGACCGGCTATTTTCC---ACCGGCTAATAGACTTCGCCGGTTCCGGCG  699
                                                |||    ||||| |||||||||
Db     61  GCCGCACGTCCGGCCGCTGCTCTTCGGGTCCGGCTCCGGGTCATCGGACTTCGCCGGCTCCGGCG  120

Qy    700  TAGTCCACATGGTCGGCGGAATAGCCGGCCTATCGGAGGCGTGATCGAGGCCCAAGAA  759
                 || |||||||||||||| |||     ||||| ||||||||| ||| ||||||
Db    121  TCGTTCCACATGGTCGGCGGCATCGCGGCGTGTGGGCGCGCTGTCGGAGGCCGCCGCA  180

Qy    760  TGGGACGTTCGATCATCGAGGACGAGCTGTGGCCTTGCGAGGCCACAGCCGGTCCTTAG  819
                |||  ||||| |||||||||| ||||||| ||||||||| |||||||| ||||||
Db    181  TCGGGGCGTTCGACCACCGCGCTCAAGGGCGCCACAGCCGGTCGCTCG  240

Qy    820  TAGTCCTGGGACCTTCTTGCTTGGTTCGGTTGGTACGGATTTAACCCCGGTTCATTTA  879
                |||||| ||| ||||| |||| |||| |||||||| ||||||| |||| |||
Db    241  TGGTGCTCGGCACCTTCCTGCTGTGGTTGGCTGTGTACGGTTCAACCCCGGTCCTTCA  300

Qy    880  ACAAAATCCTACTTACTTACGTAACTCAGGAAATTACTACGGTCAATGAGCGCGGTTG  939
                |  |||| ||| || |||||||| |||| || |||||| ||||| ||||| ||||
Db    301  CCACCATCCTCAAGTCGTACGGCCCCGGCACCGTCCACGGCAGTGTCGGCCGTGG  360

Qy    940  GCAGAACCGCGGTCACCACTACCCTAGCGGGGTCAACAGCTGCCTTGACCACGCTATTCG  999
                |||| |||||| || |||||||| |||||||||  |||| |||||||| ||||| ||
Db    361  GCCGCCACCGCGTCACCACCACCCTCGCGCCAGCGGTCGCCGCGCCCTCACCACGCTGTTCG  420

Qy   1000  GTAAACGGGTGATATCCGGTCACTGGAACGTGACCGATGTCTGCAACGGGCTGTTAGGCG  1059
                || | |||| ||||||||| ||||||| |||||||| ||||| |||||| |||||||
Db    421  GGAAGCGGCTCCAGACGGGCCACTGGAACGGGTGGACGTTCGCAACGGCCTCCTCCGGCG  480

Qy   1060  GTTTCGGGGGGATAAACAGCCGGTTGCTCCGGTTGAGCCATGGGCCATCGATCGTATGCG  1119
                |||||||| |||||| |||||||||||||| || ||| ||||| |||||||||||||
Db    481  GGTTCGGGGGCATCACGGCGCGGGTGCAGCGGTGGGCGGGCCGTCATCTGCG  540

Qy   1120  GTTTTGTTGCTTCTATAGTATTAATAGCTTGCAACAAATTAGCAGAGAAGTTAAGTTCG  1179
                |||||||||||||||  || |||   |||||||||  ||||| ||||| ||||||||
Db    541  GGTTCGTGTCCGGTGGGTGCTCATCGCGCCGCCAACGCGGTGCGCCCTCGCCGGTTCG  600

Qy   1180  ACGATCCTCTGGAGGCGGCGCAGTTGCACGTGGTGGCACGTGGGGGTGATATTCA  1239
                |||| ||| ||||| ||||||||||| ||||||| |||| ||||||| ||||||||
Db    601  ACGACCCCGCTGGAGGCGGCGGCCAGCTGCACGCGGGGTGGCGCCTGGGGGGTCCTCTTCA  660

Qy   1240  CGGCGTTGTTGCTCCAAAAAAAGGAGTATGTGAAGGAGTTTACGGTTGGGAGGAGGGGCCACG  1299
                       ||| |||||| ||||||   ||| |||||||||||| ||||| ||||
Db    661  CGGGGCTCTTCGCGAGGCGAAAGTACGTTGGAGAGATCTACGCGGGGAGGGCCCTACG  720

Qy   1300  GGTTGCTCATGGGGGTGTGTGGGAAGTTGCTCGGCGGCGCACGTGATTCAGATTCTGGTGA  1359
                ||||||||| ||||||||||||| ||||||||| |||  || ||||| || ||||||||
Db    721  GGCTGTTCATGGGCGGCGGCGGGGAAGCTCCTCCGCGGCAGATCATCCCAGATCCTGGTGA  780

Qy   1360  TTGCTGGGTGGGTTAGTGCGACCATGGGACCCCTTGTTTTGGGGTTGAATAAAACTGAAGC  1419
                ||||| || ||||| |||||| |||||||||||   ||||||||||  |||  |||||
Db    781  TCGCCGGGTGGGTGAGCTGCACCATGGGCCCGCTCCTCACGCGCTCGCTCAAGAAGCTGGCC  840

Qy   1420  TGTTGAGGATTCTTCAGAGGATGAGCTTGCGGGGATGGACATGACTCGCCATGGAGGCT  1479
                |||||| |||| |||| ||||||||| ||||| |||| ||| ||| ||||| ||||
Db    841  TGCTGCCATCTCGGCCGACGACGAGATGTCCGCGCATGAGCCATGGACGCACGGCGCGGCT  900

Qy   1480  TTGCTTATGCTTATGAGGATGATGA  1504
                ||||| ||| |||| |||| ||||
Db    901  TCGCCTACGTCTACCAGGACGAGGA  925
```

```
Qy   1006  GGGTGATATCCGGTCACTGGAACGGATGGCTCTGCAACGGGCTGTTAGGCGGTTCG  1065
                |||||||||||||||||||||||||||||| ||||||||||||||||| |||||||
Db      1  GGATGATATCCGGTCATTGGAACGTGACCGATGTCTGCAACGGGCTGTTAGGCGGCTTCG    60

Qy   1066  CGGCGGATAAACAGCCGGTTGCTCCGTGGTTGAGCCATGGCAGCCATCGTATGCCGGTTTTG  1125
                || |||||| ||||||||||| |||||| |||||||||||||||||||| ||||| ||||
Db     61  CCGCCCATAAACAGCCGGCTGCTCCGTGGTTGAGCCATGGCAGCCATCGTGTGTGGTTTTG  120

Qy   1126  TTGCTTCTATAGTATTAATAGCTTGCAACAAATTAGCAGAGAAGGTTAAGTTCGACGATC  1185
                |||||||| ||  || ||   ||||||||||||| ||||||||| ||||||||||||||
Db    121  TCGCTTCAGTGGTTCGTGATAGCCGTGCAACAAATTAGCAGAGAAGGTTAAGTTCCGATGATC  180

Qy   1186  CTCTGGAGGCGGCGGCAGTTGCACGGTGGTGTGCACGTGGGGGGGGTGATATTCACGGCGT  1245
                |||| ||| |||||||| || |||| |||||| |||||||||| || ||||||||||||
Db    181  CTTTGGAAGCGGCGGCGGCAGTTACACGGTGGTGTGCGCGGGGGGTGATATTCACGGCGC  240

Qy   1246  TGTTCGCAAAAAAGGAGTATGTGAAGGAGTTTACGGGGTTGGGGAGGGCGCACCGGTTGC  1305
                ||||||||||||| |||||||| ||||||||||| |||||| ||| ||||||||||||
Db    241  TGTTCGCGAAAAAGGAGTATGTGAAGGAGTTTATGGGGTTGGGGAGGGGAGGGCGCACGGTTGT  300

Qy   1306  TCATGGGGGGTTGGTGGGAAGTTGCTCGGCGGCGCACGTGATTCGTGATTGCTG  1365
                |||||||| |||| ||||||||||| ||||||||||| ||||||| ||| |||
Db    301  TCATGAGGGGTGGAGGGAAGTTGCTCGGCGGCGCACGTGATTTCAGATTTTGGTTATTGTTG  360

Qy   1366  GGTGGGTTAGTGCGACCATGGGACCCCTTGTTTTTGGGGTTGAATAAACTGTTGA  1425
                ||||||||||||||||||| |||||||||||||| ||||||||||  ||| |||||
Db    361  GGTGGGTGAGTGCGACCATGGGACCCTTGTTTTTGGGGTTGAATAAATTGTTGA  420

Qy   1426  GGATTCTTCAGAGGATGAGCTTGCGGGGATGGACATGAGTCGCCATGGAGGCTTTGCTT  1485
                |||||||||||||||||||||| ||||| ||||| ||| || |||||| ||||||||||
Db    421  GGATTTCTTCCGAGGATGAGCTTGCGGGGATGGATGATCTTTACCCGTCATGGAGGATTTGCTT  480

Qy   1486  ATGCTTATGAGGATGATGAGACGCACAAGCCATGGCACAAGGGTTGGCCCA  1545
                |||||||||||||||||||| |||| ||||||| |||||||||| |||||||
Db    481  ATGCTTATGAGGATGATGAGTGCCACAAGCATGGGATTCAGCTGAGGAAGGTTGGGCCCA  540

Qy   1546  ACGCGTCTTCCACACCCACTGATGAATGATCTTTTTTCCCATATGCATGTCTCATT  1605
                |||||||||||||||||||||||||||||      ||  ||||||| ||||||
Db    541  ACGCGTCGTCCACACCCACTGATGAATGATTACGATCACGATTAATTCCGGCCCCGAC  600

Qy   1606  AGTCAAACATTAAATTTGGAT  1626
                ||||||| ||||||| ||||
Db    601  AGTATTATCTTCAATTGAAAT  621
```

Query Match   27.5%; Score 469; DB 12; Length 1498;
Best Local Similarity   69.9%; Pred. No. 7.3e-114;
Matches   647; Conservative   0; Mismatches 275; Indels   3; Gaps   1;

```
;  APPLICANT: Zhou Yihua
;  APPLICANT: Cao Yongwei
;  TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53223)B
;  CURRENT APPLICATION NUMBER: US/10/424,599
;  CURRENT FILING DATE: 2003-04-28
;  NUMBER OF SEQ ID NOS: 285684
;  SEQ ID NO 109084
;  LENGTH: 459
;  TYPE: DNA
;  ORGANISM: Glycine max
;  FEATURE:
;  NAME/KEY: unsure
;  LOCATION: (1)..(459)
;  OTHER INFORMATION: unsure at all n locations
;  FEATURE:
;  OTHER INFORMATION: Clone ID: PAT_MRT3847_69516C.1
US-10-424-599-109084

Query Match             18.4%;  DB 12;  Length 459;
Best Local Similarity   84.1%;  Pred. No. 5.7e-73;
Matches  386;  Conservative  0;  Mismatches  51;  Indels  22;  Gaps  2;

Qy      469 GCTACTTCCTCTACCAATGGGGCCTTCGCCATGCGCCGCGCATCACCAGCGGAAGCA 528
                || |||||||||||||| ||||||||||||||||| || || |||||||||| |
Db        1 GCAACTTCTCTATCAATGGGGCCTTCGCCATGCGCCGCGCATCACCAGCGGCTCCA 60

Qy      529 TCGGCCGAACGGCACACAGTTCGTGGGCCTATCTCATCTACTCCTCCTTCCTCACGGCTTCG 588
            ||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db       61 TCGGCCGAACGGCACACAGTTCGTGGGCCTATCTCATCTACTCCTCCTTCCTCACGGCTTCG 120

Qy      589 TCTATCCGGTGGTCTCCCACTGGTTCGGGTCCCCAGACGGCGTGGGCCTTCTGCCTTTAAGA 648
            |||||  ||| |  ||||||||||||||||| |||| ||||||||||||  ||| |||
Db      121 TCTACCCGGTCGTCTCCCACTGGTTCGTGGTCCGCAGACGGCGTGGGCTTCTGCCATTTCCA 180

Qy      649 TCAC--------------CGACCGGCTATTTTCCACCGGCGTAATAGACTTCG 687
            ||||              ||| |||||| |||| |||||||| |||||||||||
Db      181 GCGCAGATCATGTATTCTCCCCCGGAGACCGGCTATTGTCCACCGGCGTGATAGACTTCG 240

Qy      688 CCGGTTCCGGCGTAGTCCACATGGTCGGCGGAATAGCCCGGCCTATGGGGAGGCCGCTGATCG 747
            |||||||||||||||| |||||||||||| |||| ||||||||||| ||   |||||||||
Db      241 CCGGCTCCGGCGTAGTCCACATGGTTCGGCGGTAGCGCGGCCTTCTGGGGCCGCACTGATAG 300

Qy      748 AAGGCCCAAGAATGGGACGTTTCGATCA-TGCAGGACGAGCTGTGGCCTTGCGAGGCCAC 806
            |||||  ||||||| |||||||||||||| ||||||||||| |||||||| ||||||||||
Db      301 AAGGCCCGAGAATCGGACGCTTCGACCACGGCGGAACGCGCCGNTGCCCCTCAGAGGCCAC 360

Qy      807 AGCGCGTCCTTAGTAGTCCTTGGGAACCTTCTTGCTTTGGTTCGGTTACGGATTTAAC 866
            |||||||||||||||||| ||||||||||||||| |||||||||||||||||||||||
Db      361 AGCGCGTCCTTAGTAGTCCTTGGGGACCTTCATGCTTTGGTTCGGTTACGGATTTAAC 420

Qy      867 CCCGGTTCATTTAACAAAATCCTACTTACTTACGGTAAC 905
            |||||||||||| ||| |||| |||||||||||||||||
Db      421 CCTGGTTCAATTAATAAAATCCTAGTTTCCTACGGTAAC 459
```

```
;  APPLICANT: La Rosa Thomas J
;  APPLICANT: Kovalic David K
;  APPLICANT: Zhou Yihua
;  APPLICANT: Cao Yongwei
;  TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53223)B
;  CURRENT APPLICATION NUMBER: US/10/424,599
;  CURRENT FILING DATE: 2003-04-28
;  NUMBER OF SEQ ID NOS: 285684
;  SEQ ID NO 139607
```

```
;  LENGTH: 423
;  TYPE: DNA
;  ORGANISM: Glycine max
;  FEATURE:
;  OTHER INFORMATION: Clone ID: PAT_MRT3847_97073C.1
US-10-424-599-139607

Query Match             16.8%;  DB 12;  Length 423;
Best Local Similarity   91.8%;  Pred. No. 1.7e-65;
Matches  302;  Conservative  0;  Mismatches  27;  Indels  0;  Gaps  0;

Qy       72 ACCATGTCGCTGCCTGCTTGTCCCGCGAACAACTGGCCCAACTTCTCGGCCCAAACACC 131
            ||||||||||||||||||||||||| |||||||||||||||||||||||||||| ||||
Db       95 ACCATGTCGCTGCCAGATTGTCCCGCGTCCCAACTGCCCAACTCCTGGGCCCAAATACC 154

Qy      132 ACAGACGGCCTCCGGCGCGCCGGCCTCCTTATCTGCGCCATTCGCCGCCGTGGACAGCAAG 191
            |||| |||||||||||| ||||||||||||||||| ||| ||||| |||||||||||||||
Db      155 ACAAACGCTGCCGGCGCGCCGGCCTCCTTCATCTGCGACACGGTTCACCGCCGTGGACAACAAG 214

Qy      192 TTCGTCGACACGGCCTTCGCGTCGACAACACCTACCTCCTCTTTTCGCGTCGCCAGGAACACC 251
            |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db      215 TTCGTCGACACGGCCTTCGCGGTCGCACACACTTACCCTCCTCTTCTCCGCTCACCTCGTC 274

Qy      252 TTTTCTATGCAGCTCGGCTTCGCCATGCTTCGCGGCGTCCGTCCGCGCGCCAAGAACACC 311
            ||  | |||||||||||||||||||||||| |||||||||||| |||||||||||||||||
Db      275 TTCTCTGATGCAGCTCGGCTTCGCCATGCTTCGCGGCGTCCGTCCGCGCGCCAAGAACACC 334

Qy      312 ATGAACATCATGCTCACCAACGTTCCTGACGCTCCTGACGCGCCGCGCTCTTCTACTACCTC 371
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      335 ATGAACATCATGCTCACCAACGTTCCTCACGCGCCGCGCTCTTCTACTACCTC 394

Qy      372 TTCGGCTTCGCCTTGCCTTCGCGCTCCCC 400
            ||||||||||||||||||||||||||||
Db      395 TTCGGCTTCGCCTTCGCGCTCCCC 423
```

Search completed: March 9, 2004, 16:03:09
Job time : 1300 secs

```
OM nucleic - nucleic search, using sw model

           GenCore version 5.1.6
           Copyright (c) 1993 - 2004 Compugen Ltd.

Run on:    March  9, 2004, 10:35:31 ; Search time 4638 Seconds
                                       (without alignments)
                          10984.235 Million cell updates/sec

Title:         US-10-033-109-3
Perfect score: 1706
Sequence:      1 gcacgagcactcccaacccc.........aaaaaaaaaaaaaaaaaaaa 1706

Scoring table: IDENTITY_NUC
           Gapop 10.0 , Gapext 1.0

Searched:      27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters:    55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   EST:*
        1:   em_estba:*
        2:   em_esthum:*
        3:   em_estin:*
        4:   em_estmu:*
        5:   em_estov:*
        6:   em_estpl:*
        7:   em_estro:*
        8:   em_htc:*
        9:   gb_est1:*
       10:   gb_est2:*
       11:   gb_htc:*
       12:   gb_est3:*
       13:   gb_est4:*
       14:   gb_est5:*
       15:   em_estfun:*
       16:   em_estom:*
       17:   em_gss_hum:*
       18:   em_gss_inv:*
       19:   em_gss_pln:*
       20:   em_gss_vrt:*
       21:   em_gss_fun:*
       22:   em_gss_mam:*
       23:   em_gss_mus:*
       24:   em_gss_pro:*
       25:   em_gss_rod:*
       26:   em_gss_phg:*
       27:   em_gss_vrl:*
       28:   gb_gss1:*
       29:   gb_gss2:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID		Description	
1	512.4	30.0	986	14	CK251554		CK251554	EST735191
2	490.8	28.8	795	14	CD575822		CD575822	UCRPT01_0
3	489.8	28.7	836	29	CC722552		CC722552	OGVBV66TV
4	485.6	28.5	970	14	CK263329		CK263329	EST709407

c	5	485.4	28.5	775	29	CG219364	CG219364	OGWGX30TV
	6	483.8	28.4	941	14	CK290978	CK290978	EST753692
	7	482.4	28.3	968	14	CK282541	CK282541	EST745263
	8	481.4	28.2	925	14	CK290394	CK290394	EST753108
	9	474	27.8	970	14	CK289847	CK289847	EST752569
	10	463	27.1	971	14	CK287788	CK287788	EST750510
	11	460.2	27.0	975	14	CK294612	CK294612	EST757326
	12	453.8	26.6	961	14	CK285989	CK285989	EST748711
	13	451.4	26.5	722	29	CC745155	CC745155	ZMMBb012
	14	443.2	26.0	847	10	BE998772	BE998772	EST430559
	15	437.6	25.7	984	14	CK280628	CK280628	EST743350
	16	436	25.6	927	14	CK290587	CK290587	EST753301
	17	433.2	25.4	962	14	CK295232	CK295232	EST757946
	18	429	25.1	954	14	CK294795	CK294795	EST757509
	19	428.4	25.1	815	14	CB895201	CB895201	EST647993
	20	424.8	24.9	866	14	CF207646	CF207646	CAB20002
c	21	422.2	24.7	703	29	CG263379	CG263379	OG0BT15TV
	22	421.8	24.7	949	14	CK288020	CK288020	EST750742
	23	419.4	24.6	706	14	CA781520	CA781520	022G09AF
c	24	418.8	24.5	926	14	CK282540	CK282540	EST745262
	25	418	24.5	913	14	CK286736	CK286736	EST749458
	26	417	24.4	860	14	CK282387	CK282387	EST745109
	27	416.6	24.4	858	14	CK294296	CK294296	EST757010
	28	415.4	24.3	765	12	BG645673	BG645673	EST507292
	29	415	24.3	865	14	CK298495	CK298495	EST761209
	30	411.8	24.1	796	29	CG438838	CG438838	OGVGE86TH
	31	409	24.0	777	14	CB619922	CB619922	OSIIEa04M
	32	408.6	24.0	739	29	CG220307	CG220307	OG0AW68TH
	33	407.8	23.9	909	14	CK298198	CK298198	EST760912
	34	403.8	23.7	770	14	CB891201	CB891201	EST648170
c	35	403.6	23.7	971	14	CK282386	CK282386	EST745108
	36	403.2	23.6	884	14	CK294339	CK294339	EST757053
	37	403.2	23.6	927	14	CK279062	CK279062	EST725140
	38	402.6	23.6	804	28	BZ529693	BZ529693	OGAJO29TC
	39	401.8	23.6	779	14	CB892609	CB892609	EST645401
	40	398.8	23.4	842	14	CK280629	CK280629	EST743351
	41	395.2	23.2	602	9	AU089425	AU089425	AU089425
	42	390	22.9	913	14	CK261120	CK261120	EST707198
	43	383.8	22.5	822	14	CK291607	CK291607	EST754321
c	44	377.4	22.1	853	14	CK097232	CK097232	UB44DPC08
	45	374	21.9	690	14	CD861700	CD861700	AZO1.003G

ALIGNMENTS

```
RESULT 1
CK251554
LOCUS       CK251554              986 bp    mRNA      linear   EST 12-DEC-2003
DEFINITION  EST735191 potato callus cDNA library, normalized and full-length
            Solanum tuberosum cDNA clone POCBO35 5' end, mRNA sequence.
ACCESSION   CK251554
VERSION     CK251554.1  GI:39804693
KEYWORDS    EST.
SOURCE      Solanum tuberosum (potato)
ORGANISM    Solanum tuberosum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamiids; Solanales; Solanaceae; Solanum.
REFERENCE   1  (bases 1 to 986)
AUTHORS     Buell,C.R., Hart,A., Zismann,V., Karamycheva,S.A. and Baker,B.
TITLE       Generation of ESTs from potato callus tissue
JOURNAL     Unpublished (2003)
COMMENT     Contact: Robin Buell
            The Institute for Genomic Research
            9712 Medical Center Dr, Rockville, MD 20850, USA
            Email: potato-array@tigr.org
            Clones can be requested from TIGR via potato@tigr.org
            Seq primer: ATT TAG GTG ACA CTA TAG.
FEATURES             Location/Qualifiers
     source          1..986
                     /organism="Solanum tuberosum"
                     /mol_type="mRNA"
```

```
Db    857 ||||||||||  |||||||||||| |||| |||||||  ||||||| |||||||||||||| |||||
          CGGGTTAACCCGGGTTCATTTAATAAAATCCTAATTACTTACGACGGTGCAACCGGAGG  916

Qy    914 TTACTACGGTCAATGAGCGCGGTTGGCAGAACCGGGTCACCTACCCTAGCGGGTC  973
          |||||| ||||||||||||||| |||| ||||||||| ||||||||||||||||  ||
Db    917 GTATTATGGTCAATGAGCGCGTGTAGGCCGAACCGCGGTTACAACTACCCTAGCGGGTTG  976

Qy    974 AACAGCTGCC  983
          |||| ||||
Db    977 CACCGCGCC   986
```

```
RESULT 2
CD575822
LOCUS       CD575822       795 bp   mRNA   linear   EST 12-JUN-2003
DEFINITION  UCRPT01_01af10.g3 Poncirus trifoliata CTV-challenged cDNA library -
            AGI Poncirus trifoliata cDNA clone UCRPT01_01af10, mRNA sequence.
ACCESSION   CD575822
VERSION     CD575822.1  GI:31671724
KEYWORDS    EST.
SOURCE      Poncirus trifoliata
  ORGANISM  Poncirus trifoliata
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Sapindales; Rutaceae; Poncirus.
REFERENCE   1 (bases 1 to 795)
  AUTHORS   Roose,M.L., Ye,X., Federici,C.F., Close,T.J., Fenton,R.D.,
            Wanamaker,S., Collura,K., Feuerbacher,O., Kim,H.R., Kudrna,D.,
            Wing,R. and Yu,Y.
  TITLE     Development of EST Resources and New Genetic Markers for California
            Citrus - Poncirus trifoliata CTV-challenged phloem - AGI
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Mikeal Roose
            Department of Botany & Plant Sciences, University of California
            Riverside, CA, 92521-0124, USA
            Tel: 9097874137
            Fax: 9097874437
            Email: mikeal.roose@ucr.edu
            Seq primer: T3.
FEATURES             Location/Qualifiers
     source          1..795
                     /organism="Poncirus trifoliata"
                     /mol_type="mRNA"
                     /cultivar="Pomeroy OP"
                     /db_xref="taxon:37690"
                     /clone="UCRPT01_01af10"
                     /tissue_type="Phloem"
                     /dev_stage="10 - 30 cm shoots"
                     /lab_host="E. coli TJC121"
                     /clone_lib="Poncirus trifoliata CTV-challenged cDNA
                     library - AGI"
                     /note="Vector: Lambda Uni-ZAP XR, excised phagemid;
                     Site_1: EcoRI; Site_2: XhoI; Plants were grown in the
                     greenhouse at University of California, Riverside. The
                     scion was a open-pollinated (very probably selfed)
                     seedling of Poncirus trifoliata cv Pomeroy that was
                     selected as homozygous for the Ctv resistance gene. The
                     rootstock was sweet orange infected with citrus tristeza
                     virus (CTV) isolate T514 over 1 year before sampling (CTV
                     infects sweet orange, but not genotypes carrying the Ctv
                     resistance gene. Shoots 10-30 cm long were harvested in
                     October 2000, and the green phloem (bark) was removed and
                     frozen quickly in dry ice. Total RNA was purified using
                     Trizol reagent (Gibco). Poly(A) RNA was purified, a cDNA
                     library was made, and 0.5 million primary lambda cDNA
                     clones were in vivo excised to give a population of
                     pBluescript SK(-) phagemids. All steps to this point were
                     performed in the ML Roose lab at the University of
                     California, Riverside by X. Ye. Phagemids were plated,
                     plasmid DNA purified, cDNA clones archived, and DNA
                     sequences determined bi-directionally using an ABI3730 at
                     the Arizona Genomics Institute, University of Arizona
                     (Collura, Feuerbacher, Kim, Kudrna, Wing, Yu).
```

ORIGIN

```
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="POCBO35"
/tissue_type="callus"
/lab_host="DH10B-TonA"
/clone_lib="potato callus cDNA library, normalized and
full-length"
/note="Vector: pCMVsport6.1; Site_1: EcoRI; Site_2: NotI;
supplier: RNA was isolated from Solanum tuberosum var.
Kennebec callus tissue grown on solid media."
```

```
Query Match    30.0%;   Score 512.4;   DB 14;   Length 986;
Best Local Similarity  72.0%;
Matches 698;  Conservative  0;  Mismatches 266;  Indels  6;  Gaps  2;

Qy     17 CCCCACCCGTAGTTCTACCACCTTCAGTCACGGCGTAATACACTAACCAACCCACCAT   76
          ||||  || || | ||| | ||||  | ||||  ||
Db     20 CTCCCCTCTTCTTCCTTACACCCACCCCCTCTCACCCCCACCCCCACCCCCAC        79

Qy     77 GTCGCGCTGCCTGTCCCGCCGAACAACTGGCCACTTCTCGGCCCAAACACCACAGA    136
          | |||| |||| ||  ||||| ||  || | || |||||
Db     80 CAGCCCAATGGCTTGTTCCGTCGATACACTCGCTCCGTTCCTTCGGTCCCAACACCACCAA  139

Qy    137 CGCCTCCGCCGCCGCGCCTCCCTTATCTGGGGCCATTCGCCGCCTGGACGCAAGTTCGT   196
          |||  ||| ||||| | | |||  ||  || ||  || ||
Db    140 CGCCGTCGCCGCCGCCTCATCACTCGCAACCAATTTTCCGGTGTATCGGACAGATTCCT   199

Qy    197 CGACACGGCCTTCGCCGTCGACAACACCTACCTCCTCTTTCCGCCTACCTCGTTTTTTC   256
          ||| |||| |||| | | | ||  | ||| | || ||| ||| || ||||  ||||| |
Db    200 CGATACCGGTTATGCAATCGATACCACATATCTCCTCTTCTTCAGCTTATCTCGTTTCTC  259

Qy    257 TATGCAGCTCGGCTTCGCCATGCTGCGCCCGCGTCCGTCCGGCCGGCCCAAGAACACCATGAA  316
          ||||||||||| | ||| |||||| ||| | | ||| || ||||| || |||| |||| ||||
Db    260 TATGCAGCTCGGTTCGCTATGCTTTGCGCGGCTGCGCGCGGCCGCGGTGCAAAGAATACAATGAA  319

Qy    317 CATCATGCTCACCAAGTCCTGACGCGTCGCGCGGGGCCCTCTTCTTCTACTACCTCTTCGG   376
          |||||||| || ||  ||| || |||| | ||||| || || ||  ||| |||| ||||
Db    320 CATCATGCTTACAAATGTCCTCGATGCTCGGCGGCGTGCCCGGTGGACTTTCTATTACCTATTCGG  379

Qy    377 CTTCGCCTTGCTTTCGGCTCCCCCCTCCAAGCGGCTTCATCGGTAAAACATTCTTCGGCCT   436
          ||  || | ||| ||||  ||  | | | ||| || || || |||  | ||| |||||| |
Db    380 ATTCGTTTCGGCTTTGGGGGCGGTCGGCTCCGGTCAATGATTCATCGGCCGACATTTCTTCGGGGCT  439

Qy    437 CAAGGACATCCCTTCATCCTCCTACGACTACAGCTACCTTCCTCTACCAATGGGCCTTCGC   496
          ||| | || | || | || ||  |||| ||||  ||  || | || | || |||| |||
Db    440 TAAAGAGATCCCGTCAAATTCATTCGATTACAGTAATTTCCCTCATCAATGGGCTTTCGC   499

Qy    497 CATCGGCGCCGCCGGCATCACCAGCGGAAGCATCGCCGAACGACACAGTTCGTGGCCTA   556
          ||||| || |||| ||| ||| ||  || || |||  | || || ||  || |||| |
Db    500 CATCGGCGCCGCCGGTATCACCAGCGGTTCAATCGCCGAACGCACCGACCCAATTGTGGGCGTA  559

Qy    557 TCTCATCTACTCCTCCTCCTCCCACCGGCTTCGTCCATCCGGGTGGTCTCCCACTGGTTCTG   616
          || | || || ||  | |  || |||| ||||| || |||| || || | ||| |||||||
Db    560 TTTGATTTATTCCTCTTTTCTAACGGGGTTCGTTACCCGGTTGTATCTCATCATTGGTTTTG  619

Qy    617 GTCCCCAGACGGCTGGGCCTCTGCCTTTAAGATCACCGACGGCTATTTTCCACCGGGGT    676
          |||  ||| |||| ||| ||   ||||   | ||| | |||| || |||  ||||||
Db    620 GTCCCCAGATGGGTGGGCTAGCCCCGGGTAA---TTCAAATCTTTTATTCGGATCCGGTGT  676

Qy    677 AATAGACTTCGCCGGTTCCGGGCGGTAGTCCACATGGTCGGCGGAATAGCCGGCCTATGGGG   736
          ||  |||| |||  || |||| || || || ||  |||| || ||||| |||  || ||
Db    677 TATCGATTTCGCCGGGGTCGGGGTCGGGGTGTAGTTCACATGGTAGGTGGGGTTTTATGG   736

Qy    737 AGCGGCTGATCGAAGCGCCCAAGAATGGACGTTCGATCATGCAGGACGAGCTGTGGCCTT   796
          || |||| || | |||  | || ||| ||  |||| || ||| || | | ||  || |
Db    737 AGCTTTAATTCAAGGTCGAGAATCGGGCGGTTCGATCATACGGGTCGATCCGTTGCACT   796

Qy    797 GCGAGGCCACAGCGCGTCCTAGTAGTCCTGGGAACCTTCTTGCTTTGGTTCGTTGGTA   856
          || ||||| || | ||| | || || ||  || ||| |  | ||| | |||| ||
Db    797 TCGTGGACATAGCGCGTCACTCGTGTTCTAGGCACCTTTTATTATGGTTCCGGATGGTA   856

Qy    857 CGGATTTAACCCGGGTTCATTTAACAAAATCCTACTTACTT---ACGGTAACTCAGGAAA   913
          ||  |  |||||||||||||||||   ||||||||| |||||     |||||  ||| ||
```

```
             Chromatogram files were transmitted to UC Riverside (by
             Yu), then processed at UC Riverside (by Wanamaker) using
             the HarvEST pipeline (http://harvest.ucr.edu) to remove
             vector and cloning oligo sequences and various
             contaminants, and to trim to a high quality region.
             Sequences that retained a phred 17 region of at least 100
             bases were deposited to GenBank."

ORIGIN

Query Match         28.8%;   Score 490.8;  DB 14;   Length 795;
Best Local Similarity 81.2%,  Pred. No. 1.2e-44;
Matches 570;  Conservative 0;  Mismatches 132;  Indels 0;  Gaps 0;

Qy    89  TTGTCCGCCGAACAACTGGCCCAACTTCTCGCGCCAAACACCACAGACGCCTTCCGCGC  148
          |||||||||||||||||||||||||||||||||| ||| |||||||||||||||
Db    78  TTGTTCAGCGGACCTACTCGCGCGGCTTCCTTGCCCCAACGCCCACCAAGCCCTGTCAGC 137

Qy   149  CGCCTCCCTTATCTGCGGCCATTCGCCGCGCCGTGGACAGCAAGTTCGTCGACACGGCCTT 208
          |||||||||     |||||||   ||| |||||| |||   ||||||| |||||
Db   138  AGCCTCCTTAATCTGTCGCCAACTCAGAAGCCATCAACAATAAGTTCACCGACACCGCCATA 197

Qy   209  CGCCGTCGACAACACCTACCTCCTCTTTTCCGCCTACCTCGTTTTTTCTATGCAGCTCGG  268
          |||||||||||||| |||||||||||||   |||  ||||||||||||||||||||||||
Db   198  CGCCGTCGACTCCAACCTACCCTCTCTTCTCCGGCCTACTTAGTCTTCTCCATGCAGCTCCG 257

Qy   269  CTTCGGCCATGCTCTCGCCGCGGCTCCGTCCGTCCGCCGGCCAAGAACACCATGAACATCATGCTCAC 328
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   258  CTTCGGCCATGCTCTCGCCGCGGCTCCGTCCGTCCGCCGGCCAAGAACACCATGAACATCATGCTCAC 317

Qy   329  CAAGCTCCTGAGCAGCGCCGCCGGCCGGCCTCCTACTACCTTCTCGGCTTCGGCCTTCGC  388
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   318  CAAGCTCCTGAGCAGCGCCGCCGGCCGGCCTCCTACTACCTTCTCGGATTCGGCCTTCGC 377

Qy   389  TTTCGGCTCCCCCTCCCAAAGGGCTTCATCGGTAAACATTTCTTCGGCCTCAAGGACATCCC  448
          |||| |||||| || ||||||||||||||||||||||||||||||||||||||| || |||
Db   378  CTTCGGCACCCCTTCTAAGGCTTCATCGGCCGCCACAACTTCGCCCTCAAGTCCTTTCCC 437

Qy   449  TTCATCCTCTACGACTACAGCTACTTCCTTACCAACATGGGCTTCGCCATGCCGCCGC  508
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   438  CACCTCCAACTTCGACTACAGCTACTTCCTCTACCATGGGCTTCGCCATGCCGCCG 497

Qy   509  CGGCATCACCAGCGGAGAAGCATCGCCGAACGCACAGTTCGTGGCCTATCTCATCTACTC  568
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   498  CGGCATCACCTCCGGCTCCATCGCCGAGCGCACCCCAGTTCGTCGCCTACTTAATCTACTC 557

Qy   569  CTCCTCCTCACCCGGCTTCGTCTATCCGGTGGTCTCCCACTGGTTCTCGGTCCCCAGACGG  628
          |||||||||||| || || ||||||||||||||| |||||||||||||||||||||||
Db   558  CTCCTTCCTCACCGGTTCGTCTCCTACCGTGTCTCCCATTGGTTCGTGGTCCGTCGACGG 617

Qy   629  CTGGGCCTCTGCCTTTAAGATCACCGACCGGCTATTTTCCACCGGGGTAATAGACTTCGC  688
          ||| |||||||||||||||||||||||||||||||||||||||||||||||
Db   618  CTGGGCCAGCGCCCTCCAACACCGGCCACCTCTGTTCGGCTCGGGGTCATCGACTTCGC 677

Qy   689  CGGTTCCGGCGGTAGTCCACATGGTCGGCGGAATAGCCGGCCTATGGGAGCCGCTGATCGA  748
          ||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   678  CGGCTCCGGCGTCGTCCACATGGTGGGAGGAATCGCCGGGACTTCGGGCGGTGCCTACATCGA 737

Qy   749  AGGCCCAAGAATGGACGTTTCGATCATGCAGGACGAGCGTGT 790
          ||| |||||||||||||||| ||||||||||||||||||
Db   738  AGGTCCAAGAATCGGACGGTTCGACCGGCTTCGACCGCTCCGGCCAGTCAGT 779
```

```
Query Match         28.7%;   Score 489.8;  DB 29;  Length 836;
Best Local Similarity 75.3%,  Pred. No. 1.6e-44;
Matches 624;  Conservative 0;  Mismatches 202;  Indels 3;  Gaps 1;

Qy   301  CCAAGAACACCATGAACATCATGCTCACCACCAAGCGTCCTGAGCGTCGCCGCGGCGGCTCT  360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   832  CCAAGAACACCATGAACATCATGCTGACCAACGTCTGACCAACGTGCTGGAGCGCCGGGGCGCTCT 773

Qy   361  TCTACTACCTCTTCGGCTTCGGCCTTCATCCTTCATCCTTCCTACGACTACACTCCTCT  420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   772  TCTACTACCTCTTCGGCTTCGGCCTTCGGCCTTCGGCGTCGGCCTCATCGGCA 713

Qy   421  AACATTTCTTCGGCCTCAAGGACATCCCTTCATCCTTCTACGACTACAGTGACTTCCTCT  480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   712  GGCACTTCTTCGGCCTCTCGGCGTCCAGGACCGGCTTCGACTACGACTTCTTCCTAT 653

Qy   481  ACCAATGGCCTTCGCCATGCCGCCGCCGGCGCATCACCAGCGGACGGCAGCATCGCCGAACCGCA  540
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db   652  ACCAGTGGCCTTCGCCATGCCGCCGCCGGCGGCATCACCAGCGGCTCCATCGCCGAGAGA 593

Qy   541  CACAGTTCGTGGCCTATCATCTCATCTACTCCTCCTCCTCACCGGCTTCGTCTATCCGGTGG  600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   592  CCCAGTTCGTCGTCCTACCTCCTCTGTTCGGCTTCGACGGGGTTCGTGTACCCCGTGG 533

Qy   601  TCTCCCACTGGTTCTCGTCCCCAGACGGCTGGGCCTCTGCCTTTAAGATCACCGACCGGC  660
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   532  CGTCCCACTGGTTCTCGTCCCGCGACGGGTGGGCGCGCAGTCGGGCGCCGGCCCGC 473

Qy   661  TATTTTCC--ACCGGCGTAATAGACTTCGCCGGTTCCGGCGGTAGTCCACATGGTCGGCG  717
          | |||       ||||||||||||||||||||||||||||||||||||||||
Db   472  TGCTCTTCGGGGTCCGGCGGCCATCGACTTCGCCGGCGTTCCACATGGTCCACATGGTCGGCG 413

Qy   718  GAATAGCCGGCCTATGGGAGCCGCTGATCGAAGGCCCAAGAATGGACGTTTCGATCATG 777
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   412  GCATCCGCGGGTTCTCGGGCGGCGCTCGATGAGCGGCCCCGTATCCGACGTTCGACCACG 353

Qy   778  CAGGACGAGCTGTGGCCTTGGGAGGCCACAGCGCGTCCTTAGTAGTCCTCGGGAACCTTCT 837
          ||||| |||| ||||||||||  |  ||||||||||||||||||||||||||||||||||
Db   352  CCGGCCGCTCCGTAGCCGCTCAGGGGCCACAGCCGTCGTCGCGCGTCGGCGGCCACCTTCC 293

Qy   838  TGCTTTGGTTCGGTTCGGTTACGGATTTAACCCCGGTTCATTTAACAAAATCCTACTTACTT 897
          |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   292  TGCTGTGGTTCGGCTGGTTACGGGTTACGGGGTTCAACCCCGGGTCGTTCACCACCATCCTCAAGTCGT 233
```

ORGANISM. Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

FEATURES   Location/Qualifiers
     source 1..836
             /organism="Zea mays"
             /mol_type="genomic DNA"
             /strain="B73"
             /db_xref="taxon:4577"
             /clone="ZMBMa0498L12"
             /clone_lib="ZM_0.7.1.5_KB"
             /note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb
             methylation filtered genomic DNA library"

ORIGIN

LOCUS   CC722552   836 bp   DNA   linear   GSS 23-JUN-2003
DEFINITION OGVBV66TV ZM_0.7.1.5_KB Zea mays genomic clone ZMBMa0498L12,
            genomic survey sequence.
ACCESSION   CC722552
VERSION   CC722552.1 GI:32141485
KEYWORDS   GSS.
SOURCE   Zea mays
  ORGANISM Zea mays

RESULT 3

```
Qy    898  ACGGTAACTCAGGAAATTACTACGGTCAATGGAGCGCGGTTGGCAGAACCGCGGTCACCA  957
           ||||  ||  || | |||    ||||| ||||| |||||| ||| ||| |||| |||||||
Db    232  ACGGCCCCGCGGGACCGGTCCACGGGCAGTGGTCGTCGGCCGTGGGCCGCACCGCCGTCACCA  173

Qy    958  CTACCCTAGCGCGGGGTCAACAGCTGCCTTGACCACGCTATTCGGTAAACGGGTGATATCCG  1017
           ||    ||  ||  ||  |||||| || ||  |  ||||||  ||   | |||| || |||||
Db    172  CCACCCCTCGCCGGCAGCGTCCGCGGCCTGACCACGCGTGTTCGGGAAGCGGCTCCAGGCGG  113

Qy   1018  GTCACTGGAACGTGACCGATGTCTGCAACACGGCGTGTTAGGCGGTTCGGCGCGATAACAG  1077
           ||||||||||||  ||  || |||| ||||||||||| ||||||||||| ||| || ||||||
Db    112  GCCACTGAACGTGTGGACGTCTCGCAACGGGCTCCTCGGCGCGGGTTCGGCGCCATCACGG  53

Qy   1078  CCGGTTGCTCCGTGGTTGAGCGATGGCAGCCATCGTATGCGGTTTGT  1126
           ||||  ||  || ||||||||||| ||||  |||| |||||||||||||
Db     52  CCGGGTGCAGCGCGTGGTGGAGCGTGGGGCGGCCGTCGTGTGCGGTTCGT  4
```

```
CK263329
LOCUS       CK263329                   970 bp    mRNA            linear       EST 12-DEC-2003
DEFINITION  EST709407 potato abiotic stress cDNA library Solanum tuberosum cDNA
            clone POABL79 5' end, mRNA sequence.
ACCESSION   CK263329
VERSION     CK263329.1  GI:39820307
KEYWORDS    EST.
SOURCE      Solanum tuberosum (potato)
  ORGANISM  Solanum tuberosum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamiids; Solanales; Solanaceae; Solanum.
REFERENCE   1  (bases 1 to 970)
  AUTHORS   Buell,C.R., Hart,A., Zismann,V., Karamycheva,S.A. and Baker,B.
  TITLE     Generation of ESTs from abiotic stressed potato tissue
  JOURNAL   Unpublished (2003)
COMMENT     Other ESTs: EST709408
            Contact: Robin Buell
            The Institute for Genomic Research
            9712 Medical Center Dr, Rockville, MD 20850, USA
            Email: potato-array@tigr.org
            Clones can be requested from TIGR via potato@tigr.org
            Seq primer: ATT TAG GTG ACA CTA TAG.
FEATURES             Location/Qualifiers
     source          1..970
                     /organism="Solanum tuberosum"
                     /mol_type="mRNA"
                     /cultivar="Kennebec"
                     /db_xref="taxon:4113"
                     /clone="POABL79"
                     /tissue_type="abiotic stress treated leaf and root tissue"
                     /lab_host="DH10B-TonA"
                     /clone_lib="potato abiotic stress cDNA library"
                     /note="Vector: pCMVSport6.1; Site 1: EcoRI; Site 2: NotI;
                     supplier: Solanum tuberosum var. Kennebec plants were
                     grown from cuttings on a 16hr light/8 hr dark cycle at 25
                     C for 3-4 weeks. Abiotic stress conditions were applied to
                     four separate sets of plants. Set 1 involved saturation of
                     the soil with 150 mM NaCl and tissues were harvested at
                     following application of the salt stress (leaves: 2hr,
                     6hr, 12hr, 1d, 2d, and 4d; roots:2hr, 6hr, 12hr, and 2d).
                     Set 2 were grown under the standard conditions and then
                     were water stressed by withdrawal of further watering
                     applications. Drought stressed plants were harvested after
                     cessation of watering (leaves: 3d, 5d, and 7d; roots:3d
                     and 5d). Set 3 were grown under the standard conditions
                     and then were cold stressed by placement at 4 C. Cold
                     stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
                     and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d,
                     2d. Set 4 were grown under the standard conditions and
                     then were heat stressed by placement at 35 C. Heat
                     stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
                     2d and 4d and heat-stressed roots were harvested at 6 hr,
                     12 hr, 1 d, and 4d. RNA was isolated from all tissues and
                     equal RNA from each tissue and stress was pooled to
```

construct the cDNA library. RNA sample."

ORIGIN

Query Match   28.5%;   Score 485.6; DB 14;   Length 970;
Best Local Similarity 71.5%;   Pred. No. 4.2e-44;
Matches 668; Conservative   0; Mismatches 259;   Indels   7;   Gaps   2;

```
Qy     17  CCCCACCCGTAGTTCTACCACCTTCAGTCACGGCGTAATACACTAACCAACCCACCAT  76
           |||| |||| || ||  |||||| ||  ||| ||||||||    |||| ||||||||||
Db     29  CTCCCCTTCTCTTCTTACACCACCACCCCTCTCACCCCCACCCCCACCCCCCAC  88

Qy     77  GTCGTTGCCTGCTGTCCCGGCCGAACAACTGGCCACTTCTCGGCCAAACACCACAGA  136
           ||| ||||| ||  ||||||| ||||||||||| |||||||||| ||| ||| ||||||
Db     89  CAGCCCAATGGCTTGTTCCGTCGATACACTCGCTCCGTTCCTCCGTTCCCAACACCACCAA  148

Qy    137  CGCCTCGCCGCGCGCCTCCCTTATCTGCGGGCCCCCATTCGCCGCCGTGGACAGCAAGTTCGT  196
           |||| |||| || | |||||||||  ||  ||||||  |||| || ||| ||| |||| |||||
Db    149  CGCCGTCGCCGCGCGCCGCCTCAACATCTGCAACCAATTTTCCGGTGTATCGGACAGATTCGT  208

Qy    197  CGACACGGCCTTCGCCGTGACAACACCTACCTTCCTCTTTTCCGCCACCTCGTTTTTTTC  256
           |||||| ||  |||||| || | ||| |||||  |||| ||||| || || || |||||||
Db    209  CGATACCGGTTATGCAATGCAATAACCACATATCCTCCTCTTTCAGCTTATCTCGTTTCTC  268

Qy    257  TATGCAGCTCGGCTTCGCCATGCTCTGCGCCGGCTCCGTCCGTCCGCCAAGAACACCATGAA  316
           |||||||||| ||  ||| ||||||||| ||||| |||  || || ||  ||||| || ||||
Db    269  TATGCAGCTCGGTTTCGCTATGCTTTGCGCGGGGCTCTGTCCGTCGGTCAAAGAATACAATGAA  328

Qy    317  CATCATGTCTCACCAACGTCCTGGACGCTGCGCGCCGGGCCTCTTCTACTACCTCTTCCG  376
           |||||||| ||| ||| |||| |||||| |||||| ||||||| ||  |||| |||||||||
Db    329  CATCATGCTTACAAAATGTCCCTCGATGCTGCCCGCCGGTGGACTTTCTATTACCTATTCCG  388

Qy    377  CTTGGCTTCGCTTCGGCTCCCCCTCCAAGGGCTTCATCGGTAAACATTTCTTCGGCGCT  436
           |||| ||  |  || |||| ||||| |||| |||||| ||||| ||| ||  ||||||||||
Db    389  ATTCGGTTTCGCTTGGGGCGGTCCGTCAATGATTCATCGGCGACATTTCATCGGCGGCT  448

Qy    437  CAAGGACATCCCTTCATCCTCTACGACTACAGCTACTTCCTTCCTACCAATGGGCCTTCGC  496
           |||| || |||| ||| || ||||||| ||||| |||| ||| ||  || ||| |||||| |
Db    449  TAAAGAGATCCCGTCAAATTCATTCGATTACAGTAATTTCCTCTATCAATGGGCTTTCGC  508

Qy    497  CATCGCGCCGCGCCGGTATCACCAGCGGTTCAATCGCCAGCACAGTTCGTCGGCCTA  556
           || ||||||| |||| || || || ||||||||||||| |||||||||||| ||||||
Db    509  CATCGGCGCCGCGCCGGTATCACCAGCGGTTCAATCGCCAGCACAGCCAATTTGTGGGCGTA  568

Qy    557  TCTCATCTACTCCTCCCTTCTCCACCGGCTTCGTCTATCCGGTGGTCTTCCCACTGGTTCTG  616
           |||| ||| |  |  ||||||  ||||||||  |||| || |||||  ||  || || |  |||
Db    569  TTTGATTTATTCCTCTTTTGTCTAACGGGTTTCGTTTTACCCGGTGTGTATCTCATTGGTTTTG  628

Qy    617  GTCCCCAGACGGCTGGGCCTTGCCTTAAGATCACCGACCGGCTATTTTCACCGGCGT  676
           |||| ||| ||| ||||| ||||| ||   ||||||| || || |  ||||| |||||
Db    629  GTCCCCAGATGGGTGGGCTAGCCCGGGTAA---TTCAAATCTTTTATTCGGATCCGGTGT  685

Qy    677  AATAGACTTCGCCGGTTCCGGCGTAGTCCACATGGTCGCGGAATAGCGGCCTCATGGG  736
           ||  |||||||||| || ||||||||| ||||| |||||||||| || ||||| |||||
Db    686  TATCGATTTCGCCGGTCGGTGTAGTTCACATGGAATTGCCGGGTTTTATGG  745

Qy    737  AGCGGCTGATCGAAGGCCCAAGAATGGACGTTTCGATCATGCGAGGACGAGCTGTGGCCTT  796
           ||||| |||| |||| ||| || |||| ||||||| |||||| | ||||| |||| ||||||
Db    746  AGCTTTAATTGAAGGTCCGAGAATCGGGCGGTGAGTCGATCACACGGGTGCATCCGTTGCACT  805

Qy    797  GCGAGGCCAGCGCGTCCTTAGTAGTCCTGGAACCTTCTTGCTTGGTTCGGTTGGTA  856
           ||  |||||| |||||||| ||  ||| ||||||  |||  ||| ||| |||  ||||
Db    806  TCGTGGACATAGCGCGCACTCGTGGTTCTAGGCACCTTTTATTATGGTTCGGATGGTA  865

Qy    857  CGGATTTAACCCCGGTTCATTTAACAAAATCCTACTTTAC---GGTAACTCAGGAA  912
           ||||  |||||| ||||||||| |||| ||  ||||| || |   |||||||||||
Db    866  CGGGTTTAACCCGGGTTCATTTAATAAAATCCTAATTACTTACGACGGGTGCAACCGGGAG  925

Qy    913  ATTACTACGGTCAATGGAGCGCGGTTGGCAGAAC  946
           ||||  ||  || ||||||| ||  |||| |||||
Db    926  GGTATTATGTCAATGGAGCGTGTAGGCGAAAC  959
```

RESULT 5

```
CG219364/c
LOCUS        CG219364              775 bp    DNA     linear   GSS 22-AUG-2003
DEFINITION   OGWGX30TV ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMa0575F12,
             genomic survey sequence.
ACCESSION    CG219364
VERSION      CG219364.1  GI:34119252
KEYWORDS     GSS.
SOURCE       Zea mays
  ORGANISM   Zea mays
             Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
             Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
             clade; Panicoideae; Andropogoneae; Zea.
REFERENCE    1  (bases 1 to 775)
  AUTHORS    Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
             Resnick,R.W., Nunberg,A., Robbins,D. and Lakey,N.
             Citek,R.W., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
  TITLE      Consortium for Maize Genomics
  JOURNAL    Unpublished (2002)
COMMENT      Other_GSSs: OGWGX30TH
             Contact: Cathy Whitelaw
             TIGR
             9712 Medical Center Drive, Rockville, MD 20850, USA
             Tel: 301-838-5843
             Fax: 301-838-0208
             Email: whitelaw@tigr.org
             Seq primer: TF
             Class: sheared ends.
FEATURES             Location/Qualifiers
     source          1..775
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBMa0575F12"
                     /clone_lib="ZM_0.7_1.5_KB"
                     /note="Vector: pBC5K-; Site_1: HincII; 0.7-1.5 kb
                     methylation filtered genomic DNA library"
ORIGIN

Query Match      28.5%;  Score 485.4;  DB 29;  Length 775;
Best Local Similarity   77.5%;  Pred. No. 4.9e-44;
Matches  601;  Conservative   0;  Mismatches 171;  Indels   3;  Gaps   1;

Qy   221 CACCTACCTCCTCTTTCCGCTCACTACCTCGTTTTTCTATGCAGCTCGGCTTCGCCATGCT 280
         ||||| |||| ||| ||||||||| |||||||   |||||||||||||||||||||||||||
Db   775 CACGTACCTGCTCTTCTCGGCCTACCTCCGTGTTCGCCATGCAGCTCGGCTTCGCCATGCT 716

Qy   281 CTGGCGCGGCTCCGTCGCCGGCCAAGAACACCATGAACATCATGCTCACCAAGTCCTGGA 340
         || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   715 GTGGCGCGGCTCCGTCGCCGGCCAAGAACACCATGAACATCATGCTCACCAAGTCCTGCGA 656

Qy   341 CGCTGCCGCCGCGGCGGCCTCTTCTACTACCTCCTCGGGTTCGCCTTCGGTCTCCCC 400
         ||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db   655 CGCCGCCGCCGCGGGCGGGGCCTCTTCTACTACCTCCTCGGGTTCGCCTTCGGCCACGCC 596

Qy   401 CTTCCAACGGCTTCATCGGTAAACATTTCTTCGGCCTCAAGGACATCCCTTCATCCTCCTA 460
         |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db   595 CTCCAACGGCTTCATCGGCATCGGCAGTTCTTCGGGCTCGGCCTCGAAGCACCTGCCCAGGACCGGCTT 536

Qy   461 CGACTACAGCTACTTCCTCTACCAATGGGCCTTCGGCCATCGCCGCCGGCGCCCATCACCAG 520
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   535 CGACTACGACTTCCTCTTCCTCTACCAGTGGGCCTTCGGCCTTCGCCCATCGTCATCTCTACTACTTCTCCTCCTAC 476

Qy   521 CGGAAGCATCGCCGAACGCACCACAGTTCGTGCGCCATCGTCGCGCCCGGGCATCAGTC 580
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   475 GGGGTCCATCGCCGCAGCCGGACCCCAGTTCGTCGCCACCCAGTTTCTCGTCGTTCCTCTCGAC 416

Qy   581 CGGCTTCGTCTATCCGTGGTGTCTCCCACTGGTTCGTTCCCCACTGGTCCCCAGACGGCTGGGCCTTGC 640
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   415 GGGGTTCGTTCGTCTACCCCGTGTGTCGCACTGGTTCTGTCGCCCCCGACCAGCGCGGGGCCGGGGGC 356

Qy   641 CTTTAAGATCACCGACCGGCTGTATTTCC--ACCGGCGTAATAGACTTCGCCGGTTCCGG 697
          | |   || |||  ||||    | |  || | |||  |||| |||||  | ||||||| |||||
```

```
ACCESSION    CK282541
VERSION      CK282541.1  GI:39854250
KEYWORDS     EST.
SOURCE       Nicotiana benthamiana
  ORGANISM   Nicotiana benthamiana
             Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
             Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
             asterids; lamiids; Solanales; Solanaceae; Nicotiana.
REFERENCE    1  (bases 1 to 968)
  AUTHORS    Buell,C.R., Hart,A., Zismann,V., Karamycheva,S.A., Day,B.,
             Staskawicz,B., Jin,H. and Baker,B.
  TITLE      Generation of EST sequences from Nicotiana benthamiana
  JOURNAL    Unpublished (2003)
COMMENT      Other_ESTs: EST745262
             Contact: Robin Buell
             The Institute for Genomic Research
             9712 Medical Center Dr, Rockville, MD 20850, USA
             Email: potato-array@tigr.org
             Clones can be requested from TIGR via potato@tigr.org
             Seq primer: ATT TAG GTG ACA CTA TAG.
FEATURES             Location/Qualifiers
     source          1..968
                     /organism="Nicotiana benthamiana"
                     /mol_type="mRNA"
                     /db_xref="taxon:4100"
                     /clone="NBMAD22"
                     /tissue_type="abiotic and biotic stress-treated leaves,
                     callus tissue and root tissue"
                     /lab_host="DH10B-TonA"
                     /clone_lib="Nicotiana benthamiana mixed tissue cDNA
                     library, normalized, full-length"
                     /note="Vector: pCMVsport6.1; Site_1: EcoRI; Site_2: NotI;
                     supplier: RNA was isolated from Nicotiana benthamiana
                     tissues that include callus, roots from liquid culture
                     grown plants, heat-stressed leaves (38 C, 3 hr and 6 hr),
                     cold-stressed leaves (5 C 3 hr, 6hr), and pathogen
                     challenged leaves (Pseudomonas syringae pv tomato 12 hr,
                     Xanthomonas campestris pv campestris 12 hr, 18hr;
                     Pseudomonas syringae pv phaseolicola 18hr, and Xanthomonas
                     campestris pv vesicatoria 18hr). RNA was isolated from
                     these tissues and pooled in approximately equal molar
                     amounts."
ORIGIN
```

```
normalized, full-length Nicotiana benthamiana cDNA clone NBMAD22 5'
end, mRNA sequence.
```

```
Query Match      28.3%;  Score 482.4;  DB 14;  Length 968;
Best Local Similarity  72.4%;  Pred. No. 9.4e-44;
Matches 640;  Conservative   0;  Mismatches 241;  Indels  3;  Gaps  1;

Qy      87 GCTTGTCCCGCCGAACAACTGGCCCAACTTCTCGGCCCAAACACCACCAGACGGCCTCCGCC 146
           ||||  || |||||  ||  |  |  |||  ||||  ||||||||||||| ||||||||||
Db      88 GCTTGTGAAGTGAACCAACTGGCTCCATTCCTCGGACCCAACCACCAGACGGCCGTAGCC 147

Qy     147 GCCGCCTCCCTATCTGCGGCCATTCGCCGCCGTGGACAGCAGCAAGTTCGTGACACGGCC 206
           ||||||       ||  |||  ||||||||||  ||||| |||||||| |||||||||||
Db     148 GCTGCCACTTACATATCGTAGCCAATTCTCGACATCTCCGATGTGTCCAACAAGTTTATTGATACCGGC 207

Qy     207 TTCGCGTCGACAACACCTACCTCCTCGTTTTTCTCATGCAGCTC 266
           ||||||      ||  |  ||||||  ||||||||  ||||||||
Db     208 TACGCTATCGGACTCGACTTATCTCCTCTTCCTCGGCTTACCTTGTTTTTCCATGCAGCTC 267

Qy     267 GGCTTCGCCATGCTCTGCGCCGGCGTCCGTCGCCGCGGCCAAGAACACCATGAACATCATGCTC 326
           |||||  |||||  |||||||||||||||||||||||| ||||| ||  |||||||||||||||
Db     268 GGCTTCGCTATGCTTTGTGCGGCTCTGTCCGTGGCGAAAAATACAATGAACATTATGCTT 327

Qy     327 ACCAAGTCTGGACGGTGCCCCGCCGCGGCCGCTTCATCGGTAAACATTCTTCGGCCTTC 386
           ||||||||||||||||| || | | ||||| |||||||||||  ||||| ||||||||||
Db     328 ACTAAGTCTTGACGCTCCAACGGTCTCATTGGGCGTGGGATTTTTACTATCTTTTCGCTTTC 387

Qy     387 GCTTTCGGCTCCCCCCCAACGGCTTCATCGTAAACATTCTTCGGCCTCAAGGACATC 446
           ||||| |||||||| | ||||| |||||||||||||  ||||||||||||| |||||
Db     388 GCTTGGGCGGCCGGCCGTCTAACGGCTCATTGGGCGTCACTTCTTTGGGTTAAGGAGATC 447
```

```
Qy      87 GCTTGTCCCGCCGAACAACTGGCCCAACTTCTCGGCCCAAACACCACCAGACGGCCTCCGCC 146
Db      19 GCTTGTGAAGTGAACCAACTGGCTCCATTCCTCGGACCCAACCACCAGACGGCCGTAGCC 78
Qy     147 GCCGCCTCCCTATCTGCGGCCATTCGCCGCCGTGGACAGCAGCAAGTTCGTGACACGGCC 206
Db      79 GCCGCCACTTACATATGTAGCCAATTTTCCGATGTGTCCAACAAGTTTGTTGATACCGGT 138
Qy     207 TTCGCGTCGACAACACCTACCTCCTCGTTTTTCTCATGCAGCTC 266
Db     139 TACGCTATCGGACTCGACTTATCTCCTCTTCTCGGCTTACCTTGTTTTTTCCATGCAGCTC 198
Qy     267 GGCTTCGCCATGCTCTGCGCCGGCGTCCGTCGCCGCGGCCAAGAACATCATGCTC 326
Db     199 GGCTTCGCTATGCTTTGCGCGGGCCTCGTTGTCGCGCAAGACCTTCGGCCTTAAAGATATC 258
Qy     327 ACCAAGTCTGGACGGTGCCGCCGCCGCGGCCGCTTCATCGGTAAACATTCTTCGGCCTTC 386
Db     259 ACTAAATGTTCTTGACGCCGGCGAGGTGGGATTTTTTACTACACCTCTTCCGGTTTCGGCTTTC 318
Qy     387 GCTTTCGGCTCCCCCCCAACGGCTTCATCGTAAACATTTCTTCGGCCTCCAAGGACATC 446
Db     319 GCTTGGGGGCGGGCCGGTCTAACGGCTTCATTGGGCGTCACTTCTTTGGGCTTAAAGATATC 378
Qy     447 CCTTCATCCTCCTACGACTACAGCTACTTCCTCTACCAATGGGCCTTCGCCATCGCCGCC 506
Db     379 CCTTCTAGTTCTTTTGATATACAGTAATTTCTGTACCAATGGCTTTTGCTATAGCCGCC 438
Qy     507 GCCGGCATCACCACGGCGGAAGCATCGCCGAACGCACACAGTTCGTGGCCTATCTCATCTAC 566
Db     439 GCTGGAATTACTAGCGGCTCTATAGCCGAGAGGACTCAGTTCGTGGCTTATTGATTTAC 498
Qy     567 TCCTCCTCCTCCTCCACCGGCGTTCGTCTATCCGGTCGTCCCCACTGGTTCTCGGTCCCCAGAC 626
Db     499 TCTTCTTTTCTTACCCGGCTTTGTTTACCCGGTTGTTTCTCATTGGTTTTGGGCCCCGGAT 558
Qy     627 GGCTGGGGCCTCTGCCTTTAAGATCACCGACCGGCTATTTTCCACCGGCGTAATAGACTTC 686
Db     559 GGGTGGGCCAGCCCCAACTAA---TTCAAATTTATTATTCAGGACTGGTGTTATCGACTTT 615
Qy     687 GCCGGTTCCGGCGTAGTCCACCATGGTCGGCGGAATAGCCGGCGGAATAGCCGGGGTTTTTATGGTGCTTTAATT 746
Db     616 GCCGGGTCGGGTGTCGTTCATATGGTAGGTGGTATAGCCGGGGTTTTTATGGTGCTTTAATT 675
Qy     747 GAAGGCCCAAGAAATGGACGTTTCGATCATCGCAGGACGAGCGTGCGTTGCGAGGCCAC 806
Db     676 GAAGGTCCGAGAATCGGGCCGTTCGATCATGCGGGTCGGTCGGTTCGCTACGTGGACAT 735
Qy     807 AGCGGCGTCCTTGTTAGTAGTCCTGGGAACCTTCTTGCTTTGGTTGGTACGGATTTAAC 866
Db     736 AGCGGTTCGCTTGTCGTTGTTCAGGTACCTTTTGTTGTGGTTCGGATGGTACGGGTTTAAC 795
Qy     867 CCCGGTTCATTTAACAAAATCCTACTTACTTACGGTAACTCAGGAAATTACTACGGTCAA 926
Db     796 CCCGGTTCATTTAATTAATAAGATTCTAGTGACTTATGGTACAAGTGGAGGGTATTATGGTCAA 855
Qy     927 TGGAGCGCGGTTGGCAGAACCGGCGTCACCACTACCCTAGGGGGTCAACAGCTGCCTTG 986
Db     856 TGGAGTGCTGTGGGACGTACAGCGGTGACCACCACTTAGCGGGGTGCACCGGGCGCCCT- 914
Qy     987 ACCACGCTATTCGGTAAACGGGT 1009
Db     915 ACCACGCTCTTTGGTAAGAGGAT 937
```

```
RESULT 7
LOCUS       CK282541                 968 bp    mRNA    linear    EST 15-DEC-2003
DEFINITION  EST745263 Nicotiana benthamiana mixed tissue cDNA library,
```

/clone_lib="Nicotiana benthamiana mixed tissue cDNA
library, normalized, full-length"
/note="Vector: pCMVsport6.1; Site 1: EcoRI; Site 2: NotI;
supplier: RNA was isolated from Nicotiana benthamiana
tissues that include callus, roots from liquid culture
grown plants, heat-stressed leaves (38 C, 3 hr and 6 hr),
cold-stressed leaves (5 C 3 hr, 6hr), and pathogen
challenged leaves (Pseudomonas syringae pv tomato 12 hr;
Xanthomonas campestris pv campestris 12 hr, 18hr;
Pseudomonas syringae pv phaseolicola 18hr, and Xanthomonas
campestris pv vesicatoria 18hr). RNA was isolated from
these tissues and pooled in approximately equal molar
amounts."

ORIGIN

```
Query Match             28.2%;   Score 481.4;  DB 14;  Length 925;
Best Local Similarity   72.4%;   Pred. No. 1.2e-43;
Matches 639;  Conservative   0;  Mismatches 241;  Indels   3;  Gaps   1;

Qy     87 GCTTGTCCCGCCAACACTGGCCCAACTTCTCGGCCCAAACACCACCAGACGCCTCCGCC 146
          ||||| |||| ||||||||||||||||||| |||||||| ||| |||||||| ||||||
Db     46 GCTTGTGAAGTGAACCAACTGGCTCCATTCCTCGGACCCAACACCACCGACGCCGTAGCC 105

Qy    147 GCCGGCCTCCCTTATCTGCGGCCATTTCCCGCCGTGGACGCAGCAAGTTCGTCGACACGGCC 206
          ||||| |||||| |||||||||| |||||||  ||| ||||| |||||||| ||||||||||
Db    106 GCTGCCACTTATATCTGTAGCCAATTCTCGATGTGTCCAACAAGTTTATTGATACCGGC 165

Qy    207 TTCGCCGTCGACAACACCACCTACCTCCTTTCCCGCCTACCTCGTTTTTTCTCTATGCAGCTC 266
          ||||| |||| |||||||| ||||||||||||||||||||||||||||||||||
Db    166 TACGCATCGACTCGACTTATCTCCTCTCTCGGCTTACCTTGTTTTTGTTTTTCCATGCAGCTC 225

Qy    267 GGCTTCGCCATGCTCTCGCGCCGCTCCGTCGGCGGCGGGCCTCCTTCTACTACCTCTTCGGCTTCGCCTTC 326
          ||||| |||| ||||| ||||||||||||| || ||||||||||||||||||||||
Db    226 GGCTTCGCTATGCTTTGTGCGGGCTCGTCGGTCGGTGCGAAAAATACAATGAACATTATGATGCTT 285

Qy    327 ACCAAGTCTGACGCGTGCCGCGGCTGGTGGAGATTTTTACTATCTTTTTGCTTCGCTTTC 386
          ||||| |||| ||||||||||| |||| |||||||||||||||||||||||||||||
Db    286 ACTAAGTTCTTGACGCTGACGCTGGTGGATTTTTTACTATCTTTTTGCTTCGCTTTC 345

Qy    387 GCTTTCGGCTCCCCCCTCCCAACGCGCTTCATCGTAAACATTCTTCGGCCTCAAGGACATC 446
          ||||| |||| ||||||| |||||| ||||||||||||||||||||||||||||||||
Db    346 GCTTGGGCGGCGGCCGTCTCAACGGCTTCAACGGCTTCATTGGGCGTCACTTCTTTGGGCTTAAGGAGATC 405

Qy    447 CCTTCATCCTCCTACGACTACAGCTACTTCCTCTACCGACTACATGGGCCTTCGCCATGCGGCC 506
          |||||| |||||||||||||||| |||||||||||||||||||||||||||||||||
Db    406 CCTTCTACTTCTTTGATTACAGTAATTTCTGTACCAATGGGCATTTGCTATAGCCGCC 465

Qy    507 GCCGCATCACCAGCGGAAGCATCGCCGAACGCACACAGTTCGTGGCCTATCTCATCTAC 566
          ||||| ||| ||||||||||||||||| |||||| |||| |||||||||||||||||||
Db    466 GCTGGAATTACTAGCGGCTCTATAGCCGAGAGACTCAGTTTGTCGCTTATTTGATTTAC 525

Qy    567 TCCTCCTTCCTCACCGGCTTCGTCATCCGGTGGTCTCCACTGGTTCGTCCCCCAGAC 626
          |||||| |||||||||||||||| ||| ||||||||||||||||||||||||||||
Db    526 TCTTCTTTCTTACCGGCTTGTTTTTACCGGTTGTTTTCTCATTGGTTTGGTCCCCAGAT 585

Qy    627 GGCTGGGCCTCTGCCTTAAGATACCAGACCGGCTATTTCCACGGCGGTAATAGACTTC 686
          ||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db    586 GGGTGGGCC---AGCCCGACCAATTCAAATTTGTTATTCGGGTCCGGTGTTATCGACTTT 642

Qy    687 GCCGGTTCCGGCGTAGTCCACATGGTCGGCGGAATAGCCGGCCTATGGGAGCGCTGATC 746
          ||||| |||| ||||| |||||||| ||| ||||| ||||| ||||||||||||||||||
Db    643 GCTGGGTCGGGTGGTAGTTCATATGGGTGTAGGTGGTATAGCCGGGGTTTTATGGTGCTTAATT 702

Qy    747 GAAGGCCCAAGAATGGGACGTTTCGATCATGCAGGACGAGAGCTGTGGCCTTGCGAGGCCAC 806
          ||||| ||| |||||||||||||||||||||||||||| |||||||||||||||||||
Db    703 GAAGTCCAAGAATCGGACGTTCGATCATGCGGGATCATGCGGGTCCGGTCCGTGCGCTCCGTGGGCAT 762

Qy    807 AGCGCGTTCCTAGTAGTCCTGGGAACCTTCTTGCTTCGTTGGTTCGGTTGGTACGGATTTAAC 866
          ||||| ||||||||||||||||||||||||||| ||| |||||||||||||||||||||||
Db    763 AGCGGTTCGCTTGTGTTTGGGTACCTTTTTTGTTGTTGGTTGGTTCGGTTGGTACGGGTTTAAC 822

Qy    867 CCCGGTTCATTAACAAAATCCTACTTACTTACGGTAACTCAGGAAAATTACTACGGTCAA 926
          ||||| ||||| || ||| ||| ||||||||||||||||| ||||||||||||
```

```
Db   823  CCGGGTTCCTTTAATAAGATTCTAGTTACCTATGGTACAAGTGGAGGGTATTATGGTCAA  882
          ||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Qy   927  TGGAGCGCGGTTGGCAGAACCGGCGGTCACCACCACCTAGGCGG  969
          |||||| ||||||||||||| ||||||||||||||||||||
Db   883  TGGAGTGCTGTGGGACGTACCGCGGGTGACCACCACCTAGCGG  925
```

RESULT 9

```
LOCUS       CK289847                970 bp    mRNA    linear   EST 15-DEC-2003
DEFINITION  EST752569 Nicotiana benthamiana mixed tissue cDNA library,
            normalized, full-length Nicotiana benthamiana cDNA clone NBMBQ69 5'
            end, mRNA sequence.
ACCESSION   CK289847
VERSION     CK289847.1  GI:39868739
KEYWORDS    EST.
SOURCE      Nicotiana benthamiana
  ORGANISM  Nicotiana benthamiana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamiids; Solanales; Solanaceae; Nicotiana.
REFERENCE   1  (bases 1 to 970)
  AUTHORS   Buell,C.R., Hart,A., Zismann,V., Karamycheva,S.A., Day,B.,
            Staskawicz,B., Jin,H. and Baker,B.
  TITLE     Generation of EST sequences from Nicotiana benthamiana
  JOURNAL   Unpublished (2003)
COMMENT     Other_ESTs: EST752570
            Contact: Robin Buell
            The Institute for Genomic Research
            9712 Medical Center Dr, Rockville, MD 20850, USA
            Email: potato-array@tigr.org
            Clones can be requested from TIGR via potato@tigr.org
            Seq primer: ATT TAG GTG ACA CTA TAG.
FEATURES            Location/Qualifiers
     source        1..970
                   /organism="Nicotiana benthamiana"
                   /mol_type="mRNA"
                   /db_xref="taxon:4100"
                   /clone="NBMBQ69"
                   /tissue_type="abiotic and biotic stress-treated leaves,
                   callus tissue and root tissue"
                   /lab_host="DH10B-TonA"
                   /clone_lib="Nicotiana benthamiana mixed tissue cDNA
                   library, normalized, full-length"
                   /note="Vector: pCMVSport6.1; Site_1: EcoRI; Site_2: NotI;
                   supplier: RNA was isolated from Nicotiana benthamiana
                   tissues that include callus, roots from liquid culture
                   grown plants, heat-stressed leaves (38 C, 3 hr and 6 hr),
                   cold-stressed leaves (5 C 3 hr, 6hr), and pathogen
                   challenged leaves (Pseudomonas syringae pv tomato 12 hr; 18hr;
                   Xanthomonas campestris pv campestris 12 hr, 18hr;
                   Pseudomonas syringae pv phaseolicola 18hr, and Xanthomonas
                   campestris pv vesicatoria 18hr). RNA was isolated from
                   these tissues and pooled in approximately equal molar
                   amounts."
ORIGIN
```

```
Query Match             27.8%;  Score 474;  DB 14;  Length 970;
Best Local Similarity   72.2%;  Pred. No. 7.7e-43;
Matches   631;  Conservative    0;  Mismatches 240;  Indels    3;  Gaps    1;

Qy    87  GCTTGTCCCGCCGAACAACTGGCCCAACTTCTCGGCCCAAACAACCAGACGCCTCCGCC  146
          ||||||  | || ||||||||||| ||  | |||  |||| ||||||||||   ||  ||
Db   100  GCTTGTGAAGTGAACCAACTGGCTCCATTCCTCGGACCAACACCACCGCGAGCCGTAGCC  159

Qy   147  GCCGGCCTCCCTTATCTGCGGCCATTCGCCGCCGTGGACAGCAGGCAAGTTCGTCGACACGGCC  206
          ||  ||| ||| |||| || | ||| | | | |  |  |  |  |  || ||| ||||  ||||
Db   160  GCTGCCACTTATATCTGTAGCCAATTCTCCGATGTGTCCAACAAGTTTATTGATACCCGGC  219

Qy   207  TTCGCCGTCGACAACACCTACCTCCTCTTTTCCGCCTACCTCGTTTTTTCTCATGCAGCTC  266
          |||| | ||||| || ||| | || ||  |||||||||| | |||||||  |||||||||||
Db   220  TACGCTATCGACTCGACTTATCTCCTTCTTCCGGCTTACCTTGTTTTTTCCATGCAGCTC  279
```

```
Qy   267  GGCTTCGGCCATGCTCTGCGCGGCTCCGTCCGGCGCCAAGAACACCATGAACATCATGCTC  326
          |||||||| ||||||| | |||| || |||||| ||||||||| | || |||||| ||||||
Db   280  GGCTTCGCTATGCTTGTGCGGGCTCGTGCCGTGCGGAAAAATACAATGAACATTATGCTT  339

Qy   327  ACCAACGTCCTGACGCTGCCGCCGGCCGCCTCTCTCACTACCCTTCGGCTTCGCCTTC  386
          || || || ||||||||||||||| |    | | |||| | |||| ||||||| |||||
Db   340  ACTAATGTCTTGACGCTGCCGCCGTGGGGATTTTTACATCTTTTGGCTTCGCTTC  399

Qy   387  GCTTTCGGCTCCCCTCCAAGCGCTCATCGGTAAACATTTCTTCGGCCTCAAGGACATC  446
          |||| | || || |  ||||  || || |||| |||| ||||||||  ||||||| |||
Db   400  GCTTGGGGCCGGCGGCCGTCTCAACGGCTCATTCGGGCGTCACTTCTTTGGGCCTTAAGGAGATC  459

Qy   447  CCTTCATCCTCCTACGACTACAGTACTTCCTCTACCAAGGCCCTTCGCCATCGCCGCC  506
          |||| |  |  | ||| | ||| |||| |||| |  | | |  ||| | |  |  |||||
Db   460  CCTTCTACTTCTTTTGATTACAGTAATTTTCTGTACCAATGGGCATTTGCTATTGCTTAGCCGCC  519

Qy   507  GCCGGCATCACCAGCGGGAAGCATCGCCGAACGACACAGTTCGTCGCCTATCTCATCTAC  566
          || ||  |||| | || ||   | || |||||| | || |||| | |||||  | ||| ||
Db   520  GCTGGAATTACTACGGCGGCTTCTCTATAGCCGAGAGGACTCAGTTGTCGCTTATTTGATTTAC  579

Qy   567  TCCTCCTTCCTCCACCGGCTTCGTCTCCGATCCGTGGTCTCCCACTGGTTCTCGGTCCCCAGAC  626
          || || |||| |||||||| | |||| | |||| |  |   ||| |||||    |||| ||||
Db   580  TCTTCTTTTCTTACCGGCTTGTTTCTCATTGGTTTGTCTCCCAGAT  639

Qy   627  GGCTGGGCCTCTGCTTTAAGATCACCGGCTATTTCCACCGGCGTAATAGACTTC  686
          ||||||||| || || |  ||| |||| |   |||  ||| |  |||| |||||
Db   640  GGGTGGGCC---AGCCCGACCAATTCAAATTTGTTATTGGGTCCGGTGTTATCGACTTT  696

Qy   687  GCCGGTTCCGGTAGTCCACATGGTCGGCGGAATAGCCGGCCTATGGGGAGCGGCTGATC  746
          |||||| | |||| || ||| | |  | | ||  |||| | |  |  |  || | ||| 
Db   697  GCTGGGTCGGGTGTAGTTCATATGGTATAGCCGGGTTTTATGTGCTTTAATT  756

Qy   747  GAAGGGCCCAAGAATGGGACGTTCGATCATGCAGGACGAGCGTTGCGAGGCCAC  806
          |||||||||||| | | |||  | |  ||| || |   |  |||  | ||||| |
Db   757  GAAGGTCCAAGAATCGGACGGTTCGATCATGCGGGTGCGCCTCCGTGGGGCAT  816

Qy   807  AGCGGCGTCCTTAGTAGTCCTGGGAACCTTCTTCGCTTGGTTGGTACGGATTTAAC  866
          ||||| | || || ||   | || ||| |  | | |  ||||  ||| |||| |||
Db   817  AGCGCTTCGCTTGTGGTTTGGGTACCTTTTTGTGTGTGGTTCGGTTGGTACGGGTTTAAC  876

Qy   867  CCCGGTTCATTTAACAAAATCCTACTTACTTACCGTAACTCAGGAAATTACTACGGTCAA  926
          |||| | ||   |||| || |  | | ||||||  || | |||  | ||| | | |||||
Db   877  CCGGGTTCCTTTAATAAGATTCTAGTTACCTATGGTACCATGTGGAGGGTATTATGGTCAA  936

Qy   927  TGGAGCGCGGTTGGCAGAACCGGCGGTCACCACTA  960
          |||||| ||||||||||| ||||| ||||||| |
Db   937  TGGAGTGCTGTGGGACGTACCGCGGGTGACCACCA  970
```

RESULT 10

```
LOCUS       CK287788                971 bp    mRNA    linear   EST 15-DEC-2003
DEFINITION  EST750510 Nicotiana benthamiana mixed tissue cDNA library,
            normalized, full-length Nicotiana benthamiana cDNA clone NBMBB81 5'
            end, mRNA sequence.
ACCESSION   CK287788
VERSION     CK287788.1  GI:39864664
KEYWORDS    EST.
SOURCE      Nicotiana benthamiana
  ORGANISM  Nicotiana benthamiana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamiids; Solanales; Solanaceae; Nicotiana.
REFERENCE   1  (bases 1 to 971)
  AUTHORS   Buell,C.R., Hart,A., Zismann,V., Karamycheva,S.A., Day,B.,
            Staskawicz,B., Jin,H. and Baker,B.
  TITLE     Generation of EST sequences from Nicotiana benthamiana
  JOURNAL   Unpublished (2003)
COMMENT     Other_ESTs: EST750511
            Contact: Robin Buell
            The Institute for Genomic Research
            9712 Medical Center Dr, Rockville, MD 20850, USA
            Email: potato-array@tigr.org
```

```
FEATURES
     source          Clones can be requested from TIGR via potato@tigr.org
                     Seq primer: ATT TAG GTG ACA CTA TAG.
                     Location/Qualifiers
                     1..971
                     /organism="Nicotiana benthamiana"
                     /mol_type="mRNA"
                     /db_xref="taxon:4100"
                     /clone="NBMBB81"
                     /tissue_type="abiotic and biotic stress-treated leaves,
                     callus tissue and root tissue"
                     /lab_host="DH10B-TonA"
                     /clone_lib="Nicotiana benthamiana mixed tissue cDNA
                     library, normalized, full-length"
                     /note="Vector: pCMVsport6.1; Site_1: EcoRI; Site_2: NotI;
                     supplier: RNA was isolated from Nicotiana benthamiana
                     tissues that include callus, roots from liquid culture
                     grown plants, heat-stressed leaves (38 C, 3 hr and 6 hr),
                     cold-stressed leaves (5 C 3 hr, 6hr), and pathogen
                     challenged leaves (Pseudomonas syringae pv tomato 12 hr;
                     Xanthomonas campestris pv campestris 12 hr, 18hr;
                     Pseudomonas syringae pv phaseolicola 18hr, and Xanthomonas
                     campestris pv vesicatoria 18hr). RNA was isolated from
                     these tissues and pooled in approximately equal molar
                     amounts."

ORIGIN

Query Match              27.1%;    Score 463;   DB 14;   Length 971;
Best Local Similarity    72.1%;    Pred. No. 1.2e-41;
Matches  631; Conservative     0; Mismatches  240;   Indels    4;  Gaps     2;

Qy      87  GCTTGTCCCGCGAACAACTGGCCCAACTTCTCGGCCCAAACACCACAGAGCGCCTCCGCC  146
                ||||||  ||    |||  ||||||||||||| || | |   ||||||| ||||||
Db     100  GCTTGTGAAGTGAACCAACTGGCTCCATTCCTCGGACCACCACCACCGAGCCGCTAGCC  159

Qy     147  GCCGCCTCCCTTATCTGCGGCCATTTCGCCGCCGTGGACAGCAAGTTCGTCGACACGGCC  206
                |||||||  | ||||||||||| ||     || ||| |  ||||||||| |||
Db     160  GCTGCCACTTATATCTGTAGCCAATCTCCCGATGTGTCCAACAAGTTTATTGATACCGGC  219

Qy     207  TTCGCCGTCCACAACACCTACCTCCTCTTTCCGGCTCACCTCGTTTTTCTCATGCAGCTC  266
                ||||||          |||| || |||||||||||| ||||||     ||||||||
Db     220  TACGGCTATCGACTCGACTTATCTCCTTCTCGGCTTACCTTGTTTTCCCATGCAGCTC  279

Qy     267  GGCTTCGCCCATGCTCTCGCGCGGCTCCGTCCGCCGCGCCAAGAACACCATGAACATGCTC  326
                ||||||  |||| || |||||||||||||||| |||||||| || ||| ||||||
Db     280  GGCTTCGCTATGCTTTGTGCGGGCTCTGCTCGTCCGTGCGGCTCCGTGGGAAAAAATACAATGAACATTATGCTT  339

Qy     327  ACCAACGTCTGGACGGCTGCCGCCGGCGGCCGCCTTCTACTACCTCCTTCGGCTTCGCCTTC  386
                ||||| ||||| || | ||  || |  |     |||| |  || |||||| |  ||||| |||
Db     340  ACTAATGTTCTTGACGCTGCGGCTGCGGCTGGGGATTTTTTACTATCTTTTTTGGCTTCGCTTTC  399

Qy     387  GCTTTCGGGCTCCCCCCTCCAACGGCTTCATCGGTAAACATTCTTCGGCCTCAAGGACATC  446
                ||||||||| || | ||  ||||||| || |||| ||  ||||||||||| | |
Db     400  GCTTGGGGCGGGCCGTCTAAACGGCTTCATTGGGCGTCCACTTCTTTGGGCTTAAGGAGATC  459

Qy     447  CCTTCATCCTCCTAGGACTACAGCTACTCCTCTACCAATGGGCCTTCGGCCATGCGCGCC  506
                || ||||||   | |||||| | || |||||| | | |||||||||||||||||||||
Db     460  CCTTCTACTTCTTTGATTACAGTAATTTCTGTACCAATGGGCCATTTGCTACCATAGCCGCC  519

Qy     507  GCCGGCATCACCAGCGGAAGCCAGCATCGCCGACAGAGTTCGTGGCCTATCTCATCTAC  566
                || | | | || | |||  | ||||| || | ||||| |  ||| | |   ||| ||
Db     520  GCTGGAATTACTAGCGGCTCTATAGCCGAGAGGAGGACTCAGTTTGTCGCTTATTGATTTAC  579

Qy     567  TCCTTCCTTCCTTCACCCGCTTCGTCATCCGGCGTGGTCTCCCACTGGTTCTGGTCCCAGAC  626
                | | ||||||||||||| | |||  | || ||| |||  | ||||| | ||||||||||||
Db     580  TCTTCTTTTCTTACCGGCTTTGTTTACCCGGTGTTTCTCATTGGTTTTGGTCCCCCAGAT  639

Qy     627  GGGCTGGGGCCTTCTGCCTTTAAGATCACCGACCCGGCTATTTTTCCACCGGCGTAATAGACTTC  686
                ||||||||||||    |||||    |||  | |  |||  | |||| || ||| ||||||||||
Db     640  GGGTGGGGCC---AGCCCGACCCAATTCAAATTTGTTATTCGGGTCCGTTATCGACTTT  696

Qy     687  GCCGGTTCCGGCGTAGTCCACATGGTCGGCGGAATAGCCGGCCTATGGGAGGCGCTGATC  746
                ||||||| ||| ||  | ||  ||  || |||||| || || |  ||| ||||| ||
Db     697  GCTGGGTCGGGTGAGTTCATATGGTAGGTGTATAGCCGGGGTTTTATGGGTGCTTTAATT  756
```

RESULT 11
CK294612
LOCUS   CK294612   975 bp   mRNA   linear   EST 15-DEC-2003
DEFINITION EST757326 Nicotiana benthamiana mixed tissue cDNA library,
            normalized, full-length Nicotiana benthamiana cDNA clone NBMCO68 5'
            end, mRNA sequence.
ACCESSION   CK294612
VERSION   CK294612.1 GI:39878176
KEYWORDS   EST.
SOURCE   Nicotiana benthamiana
  ORGANISM Nicotiana benthamiana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamiids; Solanales; Solanaceae; Nicotiana.
REFERENCE   1 (bases 1 to 975)
  AUTHORS   Buell,C.R., Hart,A., Zismann,V., Karamycheva,S.A., Day,B.,
            Staskawicz,B., Jin,H. and Baker,B.
  TITLE   Generation of EST sequences from Nicotiana benthamiana
  JOURNAL   Unpublished (2003)
COMMENT   Other ESTs: EST757327
            Contact: Robin Buell
            The Institute for Genomic Research
            9712 Medical Center Dr, Rockville, MD 20850, USA
            Email: potato-array@tigr.org
            Clones can be requested from TIGR via potato@tigr.org
            Seq primer: ATT TAG GTG ACA CTA TAG.
FEATURES   Location/Qualifiers
     source   1..975
                     /organism="Nicotiana benthamiana"
                     /mol_type="mRNA"
                     /db_xref="taxon:4100"
                     /clone="NBMCO68"
                     /tissue_type="abiotic and biotic stress-treated leaves,
                     callus tissue and root tissue"
                     /lab_host="DH10B-TonA"
                     /clone_lib="Nicotiana benthamiana mixed tissue cDNA
                     library, normalized, full-length"
                     /note="Vector: pCMVsport6.1; Site_1: EcoRI; Site_2: NotI;
                     supplier: RNA was isolated from Nicotiana benthamiana
                     tissues that include callus, roots from liquid culture
                     grown plants, heat-stressed leaves (38 C, 3 hr and 6 hr),
                     cold-stressed leaves (5 C 3 hr, 6hr), and pathogen
                     challenged leaves (Pseudomonas syringae pv tomato 12 hr;
                     Xanthomonas campestris pv campestris 12 hr, 18hr;
                     Pseudomonas syringae pv phaseolicola 18hr, and Xanthomonas
                     campestris pv vesicatoria 18hr). RNA was isolated from
                     these tissues and pooled in approximately equal molar
                     amounts."

ORIGIN

Query Match   27.0%;   Score 460.2;   DB 14;   Length 975;
Best Local Similarity   72.2%;   Pred. No. 2.4e-41;
Matches 627; Conservative   0; Mismatches 238;   Indels   4; Gaps   2;

```
Qy      87  GCTTGTCCCGCGAACAACTGGCCCAACTTCTCGGCCCAAACACCACAGAGCGCCTCCGCC  146
```

```
Qy     747  GAAGGCCCAAGAATGGGACGTTCGATCATGCAGGACGACGAGCTGTGCCTTGCGAGGCCAC  806
                |||||||||||||||||||||||||||||||| ||||||||||| || |||||||||||
Db     757  GAAGGTCCAAGAATCGGACGTTCGATCATGCGGGTCGTCGGTCCGTCCGTGCGCCGTGGGCAT  816

Qy     807  AGGCGGTCCTTAGTAGTCCTGGGAACCTTCTTGCTTCTTGGTTCGGTTGGTTACGGGATTAAC  866
                |||||||| || ||||||| |||||||||||| || ||||| ||||||||| ||||| |||
Db     817  AGGCGCTTCGCTTGTGTTTTGGTACCTTTTTTTGTGTGGTTCGGTTCGGTTACGGGTTTAAC  876

Qy     867  CC-CGGTTCATTTAACAAAATCCTACTTACTTACGGTAACTCAGGAAAATACTACGGTCA  925
                || |||||   |||||||| |||| |||||||| ||||| ||||| || ||||||||||
Db     877  CCGGGGTTCCTTTAATAGATTCTAGTTTACCTATCTAGTACAAGTGGAGGGGTATTATGGTCA  936

Qy     926  ATGGAGGCGGTTGGCAGAACCGCGGTCACCACTA  960
                |||| ||||| ||||| |||| ||||||||||||
Db     937  ATGGAGTGTGTGGGACGTACCCGCGGTGACCACCA  971
```

REFERENCE   asterids; lamiids; Solanales; Solanaceae; Nicotiana.
            1 (bases 1 to 961)
AUTHORS   Buell,C.R., Hart,A., Zismann,V., Karamycheva,S.A., Day,B.,
            Staskawicz,B., Jin,H. and Baker,B.
TITLE   Generation of EST sequences from Nicotiana benthamiana
JOURNAL   Unpublished (2003)
COMMENT   Other_ESTs: EST748712
            Contact: Robin Buell
            The Institute for Genomic Research
            9712 Medical Center Dr, Rockville, MD 20850, USA
            Email: potato-array@tigr.org
            Clones can be requested from TIGR via potato@tigr.org
            Seq primer: ATT TAG GTG ACA CTA TAG.

FEATURES   Location/Qualifiers
source   1..961
            /organism="Nicotiana benthamiana"
            /mol_type="mRNA"
            /db_xref="taxon:4100"
            /clone="NBMAY92"
            /tissue_type="abiotic and biotic stress-treated leaves,
            callus tissue and root tissue"
            /clone_lib="Nicotiana benthamiana mixed tissue cDNA
            library, normalized, full-length"
            /note="Vector: pCMVSport6.1; Site 1: EcoRI; Site 2: NotI;
            supplier: RNA was isolated from Nicotiana benthamiana
            tissues that include callus, roots from liquid culture
            grown plants, heat-stressed leaves (38 C, 3 hr and 6 hr),
            cold-stressed leaves (5 C 3 hr, 6hr), and pathogen
            challenged leaves (Pseudomonas syringae pv tomato 12 hr;
            Xanthomonas campestris pv campestris 12 hr, 18hr;
            Pseudomonas syringae pv phaseolicola 18hr, and Xanthomonas
            campestris pv vesicatoria 18hr). RNA was isolated from
            these tissues and pooled in approximately equal molar
            amounts."

ORIGIN
Query Match   26.6%;   Score 453.8;   DB 14;   Length 961;
Best Local Similarity   71.7%;   Pred. No. 1.2e-40;
Matches 623;   Conservative   0;   Mismatches 242;   Indels   4;   Gaps   2;

Qy   87 GCTTGTCCCGCCGAACAACTGGCCCAACTTCTCGGCCCAAAACACCACGACGCCTCCGCC 146
          |||| |||||||| |||||||||| |||||||| || ||||||||||||||||||
Db   97 GCTTGTGAAGTGAACCAACTGGCCTCCATTCCTCGGGCCGTCAGCCCTTAAGGAGATC 156

Qy   147 GCCGGCTCCCTTATCTGCGGGCCATTCGCGCCGTGGACAGCAAGTTCGTGACACGGCC 206
          | |||| ||||||| |||||| |||||||||||||| ||||||||||||||||
Db   157 GCTGCCACTTATATCTGTAGCCAATTCTCCGATGTGTCCAACAAGTTTATTGATACGGGC 216

Qy   207 TTCGCCGTCGACAACACCTACCTCCTCTTTTCCGCCTACCTCGTTTTTTTCTATGCAGCTC 266
          || || |||| || || |||||||| ||||| ||| || |||||| ||| ||||||||||
Db   217 TTCGGGCTGCAACACACCTACCTCCTCTTTCCGGCTCGGCTTACCTCGTTGTTTTTCCATGCAGCTC 276

Qy   267 GGCTTCGCCATGCTCGCGCCGCGGCTCCGCGCCGCGGCCTCTTCCACTCTCTCGGTCCCCAGAC 326
          ||||||||||||||| ||||||||||||||| || |||| ||||||||||| |||||||
Db   277 GGCTTCGCTATGCTTTGTGCCGCGTCGCCGCGGCCGCGGCCTTCTCACTACCTCTTCTACTCCTTCCGCCTTCGGCCTCAAGAACATCATGCTC 336

Qy   327 ACCAACGTCCTGGACGCTGCCGCCGGCGGCCTCTTCTACTACCTCTTCGGCCTCAAGGACATC 386
          |||||||||| || |||||||||||||||||||||||||| |||||||||||||||||||||
Db   337 ACTAATGTTCTTGACGTCGCGCTGTGGGATTTTTACTATCTTTTTCGGCTTCGCTTTC 396

Qy   387 GCTTTCGGCTCCCCTCCAACGGCTTCATCGGTAAACATTTCTTCGGCCTCAAGGACATC 446
          ||||||| ||||| |||||| || |||||||||||||| ||||||||||||||||
Db   397 GCTTGGGGCGGGGCCCGTCTAAGGCTTCATTGGGCGTCACTTCTTTTGGGCTTAAGGAGATC 456

Qy   447 CCTTCATCCTCTACGACTACAGCTACTTCCTCTACCAAATGGGCCTTCGCCATGCCGCC 506
          ||| ||||||||||||||||||| ||| ||||| ||| |||||||||||||||| ||||
Db   457 CCTTCTACTTCTTTTGATTACATGTATTTTCTGTACCAATGGGCATTTGCTATAGCCCGCC 516

Qy   507 GCCGGCATCACCAGCCGGAAGCATCGCCGAACGAACGCACCACAGTTCGTCGTGGCCTATCTCCATCTAC 566
          ||| ||||||| |||| |||| || | ||||||||| |||||| ||||||||||||||| | |||
Db   517 GCTGAATTACTAGCGGGCTCTATAGCCGAGAGACTCAGTTTGTCGCTTATTGATTTAC 576

Qy   111 GCTTGTGAAGTGAACCAACTGGCTCCATTCCTCGGACCACCACGCCGTAGCC 170
          |||||||

Qy   147 GCCGGCTCCCTTATCTGCGGGCCATTCGCGCCGTGGACAGCAAGTTCGTGACACGGCC 206

Qy   171 GCTGCCACTTATATCTGTAGCCAATTCTCCGATGTGTCCAACAAGTTTATTGATACCGGC 230

Qy   207 TTCGCCGTCGACAACACCTACCTCCTCTTTTCCGCCTACCTCGTTTTTTTCTATGCAGCTC 266

Qy   231 TACGCTATCGACTCGACTTATCTCCTCTTCTCGGCTTACCTCGTTGTTTTTCCATGCAGCTC 290

Qy   267 GGCTTCGCCATGCTCGCGCCGCGGCTCCGCGCCGCGGCCTCTTCCACTCTCTCGGTCCCCAGAC 326

Qy   291 GGCTTCGCTATGCTTTGTGCCGCGTCGCCGCGGCCGCGGCCTTCTCACTACCTCTTCTACTCCTTCCGCCTTCGGCCTCAAGAACATCATGCTC 350

Qy   327 ACCAACGTCCTGGACGCTGCCGCCGGCGGCCTCTTCTACTACCTCTTCGGCCTCAAGGACATC 386

Qy   351 ACTAATGTTCTTGACGTCGCGCTGTGGGATTTTTACTATCTTTTTCGGCTTCGCTTTC 410

Qy   387 GCTTTCGGCTCCCCTCCAACGGCTTCATCGGTAAACATTTCTTCGGCCTCAAGGACATC 446

Qy   411 GCTTGGGGCGGGGCCCGTCTAACGGCTTCATTGGGCGTCACTTCTTTTGGGCTTAAGGAGATC 470

Qy   447 CCTTCATCCTCTACGACTACAGCTACTTCCTCTACCAAATGGGCCTTCGCCATGCCGCC 506

Qy   471 CCTTCTACTTCTTTTGATTACATGTATTTTCTGTACCAATGGGCATTTGCTATAGCCCGCC 530

Qy   507 GCCGGCATCACCAGCCGGAAGCATCGCCGAACGAACGCACCACAGTTCGTCGTGGCCTATCTCCATCTAC 566

Qy   531 GCTGAATTACTAGCGGGCTCTATAGCCGAGAGACTCAGTTTGTCGCTTATTGATTTAC 590

Qy   567 TCCTCCTTCCTCACCGGCTTCGTCATCCGGTGCTCCCACTGGTCTCGGTCCCCAGAC 626

Qy   591 TCTTCTTTTCTTACCCGCTTTGTTTACCCGGTTGTTTCTCATTGGTTTGGTCCCAGAT 650

Qy   627 GGCTGGCGCCTGCCTTAAGATCACCGACCGGCTATTTTCCACCGGCTAATAGACTTC 686

Qy   651 GGGTGGGCC---AGCCCGACCAATTCAAAATTGTTATTCGGGTCCGGTGTATCGACTTT 707

Qy   687 GCCGGTTCCGGCGTAGTCCACATGGTCGGCGGAATAGCCGGCCTATGGGAGGCCTGATC 746

Qy   708 GCTGGGTCGGGTGAGTTCATATGGTGATAGTGTATAGCCGGGTTTTATGGTGCTTTAATT 767

Qy   747 GAAGGCCCAAGAATGGGACGTTTCGATCATGCAGACGAGCTGTGGCCTTGCGAGGCCAC 806

Qy   768 GAAGTCCAAGAATCGGACGGTTGATCATGCGGGTCGGTCCGTTGCGCTCCGTGGGCAT 827

Qy   807 AGCGGCGTCCTTAGTAGTCCTGGGAACCTTCTTGCTTGGTTGGTTGGTTGGTTACGGATTTAAC 866

Qy   828 AGCGGCTTCGCTTGTGGTTTTGGGTACCTTTTTGTTGTGGTTCGGTTCGGTTGGTACGGGTTTAAC 887

Qy   867 CCCGGTTCATTTAACAAAATCCTACTTACTTACGGTAACTCAGGAAATTACTACGGTCAA 926

Qy   888 CCCGGTTC-TTTAATAAGATTCTAGTTACCTATGGTGGAGGGTATTATGGTCAA 946

Qy   927 TGGAGCGCGGTTGGCAGAACCGCGGTCAC 955

Qy   947 TGGAGTGCGTGTGGGACGTACCGCGGTGAC 975

RESULT 12
CK285989
LOCUS   CK285989   961 bp   mRNA   linear   EST   15-DEC-2003
DEFINITION   EST748711 Nicotiana benthamiana mixed tissue cDNA library,
            normalized, full-length Nicotiana benthamiana cDNA clone NBMAY92 5'
            end, mRNA sequence.
ACCESSION   CK285989
VERSION   CK285989.1   GI:39861109
KEYWORDS   EST.
SOURCE   Nicotiana benthamiana
ORGANISM   Nicotiana benthamiana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

ORIGIN

Query Match　26.5%;　Score 451.4;　DB 29;　Length 722;
Best Local Similarity　77.7%;　Pred. No. 2.6e-40;
Matches 558;　Conservative　0;　Mismatches 157;　Indels　3;　Gaps　1;

```
Qy  222 ACCTACCTCCTCTTTCCGCTACCTCGTTTTTCATGCAGCTCGGCCATGCTC 281
            |||||| ||||| ||||||||||||||  |||||||||||||||||||
Db    5 ACGNACCTGCTCTCTTCGCGGCCTACCTCGTTCTCGGCCATGCAGCTCGGCCATGCTC 64

Qy  282 TGGCGCCGGTCCGTCCGCGGCCAAGAACACCATGAACATCATGAACATCATGAACATCATGGCCTCCTGGAC 341
            ||||||||||||||||||||||||||||||||||||||| ||  |||||||||||||
Db   65 TGGCGGCCGGTCCGTCCGCGGCCAAGAACACCATGAACATCATGAACATCATCACCAAGTGCTCGAC 124

Qy  342 GCTGCCGCCGGCGGCCTCTTCTACTACCTCTTCGGCTTCGGCTTTCGGCTCCCCC 401
            ||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db  125 GCCGCCGCCGGCGGCGGCGCGTCTTCTACTACCTCTTCGGCCTTCGGCTTCGGCTACGGGACCCCG 184

Qy  402 TCCAACGGCTCATCGTGAAACATTTCTTCGGCGCCTCAAGGACAATCCCTTCATCCTCCTAC 461
            ||| ||||||||||||||||||||||  || |||||||||||||||||||||||||
Db  185 TCCAACGGCTTCATCGGCAAGCACTTCTTCGGCCTTCGGCCTCCCAGGTCGGGTTC 244

Qy  462 GACTACAGCTACTTCCTCTACCAATGGGCCTTCGGCATCGCGCGCGGCATCACCAGC 521
            ||||| ||||||||||||||||  || |||||||||||||||||||||||||||||
Db  245 GACTACGACTTCTTCCTCTTCCAGTGAGCTTTGCCATCGCCGCCGGGATCACGTCC 304

Qy  522 GGAAGCATCGCCGAACGCACACAGTTCGTGGCCTATCTCATCATCATCCTCCTTCCTCACC 581
            ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  305 GGCTCCATCGCCGAGCGCACGCACGCAGTTCGTGGCGTTACCTCACTCATCCTCCGCCTTCCTCACC 364

Qy  582 GGCTTCGTCTATCCGGTGCGTGTCCCACTGGTTCGGTCCCCAGACGCGGCTGGGCCCTGCC 641
            |||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Db  365 GGCTTCGTCGTACCCGGTGGTGTCCCACTGGCTCGTCGGTCGTGTCCGCCGACGGCCTGGGCCTGCGCCG 424

Qy  642 TTTAAGATCACCG---ACCGGCTATTTTCCACCGGCTAATAGACTTCGCCGGTTCGGC 698
            | |||||||||||    |||||||||||||||||||||||||||| |||||||||||
Db  425 TCACGGACGTCGGGGAGCTCCTTCTTCGGCCGCATCATCGGACTTCGCCGGCGGTCCAGC 484

Qy  699 GTAGTCCACATGTCGGCGGAATAGCCGGCCTATGGGAGCCGCTGATCGAAGGGCCAAGA 758
            |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db  485 GTTGTCCACATGTGGGCGGAATCGCCGGCCTCTGGGGGCCCTTCATCGAGAGGGCCCCGC 544

Qy  759 ATGGGACGTTCGATCATCGACGAGGACGAGCTGTGCCTTGCGAGCCACAGCGGCGGTTCTTA 818
            |||||||||||| | || |||||||||| || ||||||||||||||||||||||||||||||
Db  545 ATTGGCCGGTTCGACCACGCCGGCGGCGCGCCGCTCGGTGGCGCTGCGCGGCCCACAGCGGCGTGCGCTC 604

Qy  819 GTAGTCCTGGGAACCTTCTTGCTTGGTTCGGTTACGGATTTAACCCCGGTTCATTT 878
            |||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db  605 GTCGTGCTCGGCACTTTCCTGCTGTGTGGTTCGGTTCAACCCCGGGGTCGTTC 664

Qy  879 AACAAAATCCTACTTACTTACGGTAACTCAGGAAATTACTACGGTCAATGGAGCGCGG 936
            |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  665 CTCACCATCCTCAAGAGCTACGGCCCCGGCCGGCAGCATCCACGGGCAGTGGTCGGCCG 722
```

RESULT 13
CC745155
LOCUS　　　　　CC745155　　　　　722 bp　　　DNA　　　linear　　　GSS 25-JUN-2003
DEFINITION　ZMMBBb0121A09.r ZMMBb Zea mays subsp. mays genomic clone
　　　　　　　ZMMBb0121A09 3', genomic survey sequence.
ACCESSION　　CC745155
VERSION　　　CC745155.1　GI:3219762
KEYWORDS　　　GSS.
SOURCE　　　　Zea mays subsp. mays (maize)
ORGANISM　　　Zea mays subsp. mays
　　　　　　　Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
　　　　　　　Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
　　　　　　　clade; Panicoideae; Andropogoneae; Zea.
REFERENCE　　1　(bases 1 to 722)
AUTHORS　　　Yu,Y., Kim,H.R., Hatfield,J., Soderlund,C., Bharti,A.K., Messing,J.
　　　　　　　and Wing,R.
TITLE　　　　Sequencing of the maize genome
JOURNAL　　　Unpublished (2003)
COMMENT　　　Contact: Rod Wing
　　　　　　　Arizona Genomics Institute
　　　　　　　University of Arizona
　　　　　　　Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
　　　　　　　85721-0088, USA
　　　　　　　Tel: 520 626 3967
　　　　　　　Fax: 520 621 9288
　　　　　　　Email: http://genome.arizona.edu
　　　　　　　PCR PRimers
　　　　　　　FORWARD: T7
　　　　　　　BACKWARD: M13r
　　　　　　　Plate: 0121　row: A　column: 09
　　　　　　　Seq primer: M13r
　　　　　　　Class: BAC ends.
FEATURES　　　　　　　　　Location/Qualifiers
　　source　　　　　　　1..722
　　　　　　　　　　　　　/organism="Zea mays subsp. mays"
　　　　　　　　　　　　　/mol_type="genomic DNA"
　　　　　　　　　　　　　/cultivar="B73"
　　　　　　　　　　　　　/sub_species="mays"
　　　　　　　　　　　　　/db_xref="taxon:4578"
　　　　　　　　　　　　　/clone="ZMMBb0121A09"
　　　　　　　　　　　　　/lab_host="DH10B"
　　　　　　　　　　　　　/clone_lib="ZMMBb"
　　　　　　　　　　　　　/note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
　　　　　　　　　　　　　HindIII; Zea mays L. ssp. mays"

RESULT 14
BE998772
LOCUS　　　　　BE998772　　　　　847 bp　　　mRNA　　　linear　　　EST 06-OCT-2000
DEFINITION　EST430559 GVSN Medicago truncatula cDNA clone pGVSN-13017, mRNA
　　　　　　　sequence.
ACCESSION　　BE998772
VERSION　　　BE998772.1　GI:10699112
KEYWORDS　　　EST.
SOURCE　　　　Medicago truncatula (barrel medic)
ORGANISM　　　Medicago truncatula
　　　　　　　Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
　　　　　　　Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
　　　　　　　rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
　　　　　　　Medicago.
REFERENCE　　1　(bases 1 to 847)
AUTHORS　　　Fedorova,M., Pierson,B.L., Samac,D.A., Gantt,J.S., Vance,C.P.,
　　　　　　　Town,C.D., Bowman,C.L., Craven,M.B., Cho,J. and Fraser,C.M.
TITLE　　　　ESTs from senescent nodules of Medicago truncatula

```
Qy    522 GGAAGCATCGCCGAACGCACACAGTTCGTGGCCTATCCTCATCCTCCTTCCTCACC 581
          ||||  |||||||||||  |||||||||||||||||||||||||||||| |||||
Db    487 GGCTCTATAGCGGACGAACTCAGTTCGTTGTTATCTTATATACTCATCTTTTCTCACA 546

Qy    582 GGCTTCGTCTATCCGGTGGTCTCCCACTGGTTCTTGGTCCCCAGACGGCTGGGCCCTGCC 641
          ||||||||  |||||||||  ||||  ||||||||||||||||||  |||||||||||||
Db    547 GGGTTTGTTTACCCTGGTTTCTCATTGGTTCTCGGTCCGGTGATGGTGGGCTAGCGGCA 606

Qy    642 TTTAAGATCACCGACCGGCTATTTTCCACCGGCGTAATAGACTTCGCCGGTTCCGGCGTA 701
          |||||||||||||||  |  ||| ||||||||||||||||||||||||||||||||||
Db    607 ACCAACACCGGCAACCTACTCTTGGTACAGGAGTCATTGATTTTGCTCGTCTCCGGTGTT 666

Qy    702 GTCCACAATGGTCGGCGGAATAGCCGGCCTATGGGAGGCGTGATCGAGGCCCAAGAATG 761
          ||| |||||||| || |||||  ||| ||  ||| |||| |||||||| |||| |||||
Db    667 GTTCATATGGTTGGTGGTATTGCTGGTGTCTTTGGGAGCACTCATTGAGGTCACGAATT 726

Qy    762 GGACGTTCGATCATGCAGGACGAGCTGGCCTTGCGAGGCCACAGCCGTCCTTAGTA 821
          ||||| |  ||||  |||  |||  |||||||||| ||||  ||||| ||||||
Db    727 GGACGGTTCGATCACTCTGGTCGAGCCGTAGCTTCCCGGACACAGTGCTTCCTTGTT 786

Qy    822 GTCCTGGGAACCTTCTTGCTTGGTTCGGTTGGTACCGGATTTAACCCGGTTCATTTAAC 881
          ||||  || |||| ||| ||| |||| ||||| |||  |||||||||| ||| |||||
Db    787 GTTCTTGGTACCCTTCATGCTTGGGTTGGTTGGTATGTTTCAACCCTGGTTCATTCAC 846
```

RESULT 15

```
CK280628
LOCUS         CK280628              984 bp     mRNA     linear     EST 15-DEC-2003
DEFINITION    EST743350 Nicotiana benthamiana mixed tissue cDNA library,
              normalized, full-length Nicotiana benthamiana cDNA clone NBMA146 5'
              end, mRNA sequence.
ACCESSION     CK280628
VERSION       CK280628.1  GI:39950448
KEYWORDS      EST.
SOURCE        Nicotiana benthamiana
ORGANISM      Nicotiana benthamiana
              Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
              asterids; lamiids; Solanales; Solanaceae; Nicotiana.
REFERENCE     1  (bases 1 to 984)
AUTHORS       Buell,C.R., Hart,A., Zismann,V., Karamycheva,S.A., Day,B.,
              Staskawicz,B., Jin,H. and Baker,B.
TITLE         Generation of EST sequences from Nicotiana benthamiana
JOURNAL       Unpublished (2003)
COMMENT       Other_ESTs: EST743349 EST743351 EST743352
              Contact: Robin Buell
              The Institute for Genomic Research
              9712 Medical Center Dr, Rockville, MD 20850, USA
              Email: potato-array@tigr.org
              Clones can be requested from TIGR via potato@tigr.org
              Seq primer: ATT TAG GTG ACA CTA TAG.
FEATURES             Location/Qualifiers
     source          1..984
                     /organism="Nicotiana benthamiana"
                     /mol_type="mRNA"
                     /db_xref="taxon:4100"
                     /clone="NBMA146"
                     /tissue_type="abiotic and biotic stress-treated leaves,
                     callus tissue and root tissue"
                     /lab_host="DH10B-TonA"
                     /clone_lib="Nicotiana benthamiana mixed tissue cDNA
                     library, normalized, full-length"
                     /note="Vector: pCMVSport6.1; Site_1: EcoRI; Site_2: NotI;
                     supplier: RNA was isolated from Nicotiana benthamiana
                     tissues that include callus, roots from liquid culture
                     grown plants, heat-stressed leaves (38 C, 3 hr and 6 hr),
                     cold-stressed leaves (5 C 3 hr, 6hr), and pathogen
                     challenged leaves (Pseudomonas syringae pv tomato 12 hr,
                     Xanthomonas campestris pv campestris 12 hr, 18hr;
                     Pseudomonas syringae pv phaseolicola 18hr, and Xanthomonas
                     campestris pv vesicatoria 18hr). RNA was isolated from
                     these tissues and pooled in approximately equal molar
                     amounts."
```

```
FEATURES             Location/Qualifiers
     source          1..847
                     /organism="Medicago truncatula"
                     /mol_type="mRNA"
                     /cultivar="genotype A17"
                     /db_xref="taxon:3880"
                     /clone="pGVSN-13017"
                     /tissue_type="senescent root nodules"
                     /dev_stage="mixture of effective nodules from 40 day old
                     plants harvested 36 hours post shoot removal and nodules
                     collected from 2 month old plants at mid-pod stage"
                     /lab_host="E. coli strain SOLR"
                     /clone_lib="GVSN"
                     /note="Vector: pBluescript SK +/-; Site_1: EcoRI; Site_2:
                     XhoI; cDNA was prepared from polyA+ enriched RNA from the
                     mixture of effective nodules of 40 day old plants
                     harvested 36 hours post shoot removal and nodules
                     collected from 2 month old plants at mid-pod stage. The
                     cDNA was directionally ligated into the Uni-ZAP XR vector
                     from Stratagene and packaged using Gigapack III Gold
                     packaging extracts. Plasmids containing cDNA inserts were
                     excised from the recombinant lambda-ZAP phage using
                     Ex-Assist helper phage and propagated in SOLR cells."
```

ORIGIN

```
Query_Match          26.0%;    Score 443.2; DB 10;   Length 847;
Best_Local_Similarity 70.5%;   Pred. No. 1.9e-39;
Matches  592;  Conservative  0;  Mismatches 248;  Indels  0;  Gaps  0;

Qy     42 TCAGTCACGGCGTATACACTAACCAACCCACCATGTCGTCGCTTGCTCGTCCCGCCGAA 101
          ||||| |||  |||||| ||||| || || |||||||| || ||| |||||||||
Db      7 TAAGCAGCTTCGTCACCCACTACCCACCCAAACATGTCGCTCGCGCCGGAATGTTCAGTTTCG 66

Qy    102 CAACTGGCCCAACTTCTCGCCCCAAACACCACCAGACGCCTCCGCGCGCCTCCCTTATC 161
          ||| || |||| |||| || |||||  |||  ||||  ||| || ||||  |||||||
Db     67 CAGCTAGCACAGTTCATCGGCCACCGCCACCACCGCCACCGCCGCCGCCGCAAATTTCATC 126

Qy    162 TGCGGCCATTTCGCGCCGTGGACAGCAAGTCGTCGACACGGCCTTCGCCGTCGACAAC 221
          |||||||| |||| ||| ||  |||| ||||  ||  |||| | |||  |||| |||
Db    127 TGCGACCAGTTGACAGCAGTGTCAAACAGATTCGTTGACACAGATTCGTTGACACAAAC 186

Qy    222 ACCTACCTCCTCTTTTCCGCCTACCTCGGTTTTTCTATGAGAGCTCCGGCTTCGGCCATGCTC 281
          |||||||||||||| ||||| ||||||| |||| |  ||  ||| ||  || ||||| ||||
Db    187 ACCTATCTTTATTCTCAGTTCAGCTTACCTTGTCTTCTCTCATGCAACTTGGTTTTTGCCATGCTC 246

Qy    282 TGGGCCGGCTCCGTCCGGCCCAAGACCATGATGAACATGCAGGTCCGGAC 341
          ||||||||||| || |||| ||||| |||||||| || |  ||||| ||||
Db    247 TGTGCTGGTTCAGTTCGTCGCTAAGAACACACATGAACATCATGCTCACAAATGTTCTCCGAC 306

Qy    342 GCTGCCGCCGGCGGCCTCTTCTACTACCTCTTCGGCTTCGGCTTCGGCTCCCCC 401
          ||| |||||  ||||  |  |  || |||||  || | |||| || |||||||||
Db    307 GCCCGCTGCCGGCGTGTTTCCTTCACTATCTTCGGTTTTGCCTTTGCTTCGGAACACCT 366

Qy    402 TCCAACGGCTTCATCGGTAAACATTTCTTCGGGCCTCAAGGACATCCCTTCATCCTCCTAC 461
          |||||  || |||  ||||||||||  |  |  | || |  ||||| ||||| ||| |
Db    367 TCAAATGGTTCATAGGAAAACATTTTCTTCGGTCGAAGGATGTTCCGAGACAGCAAATTT 426

Qy    462 GACTACAGCTACTCCCTCACCAATGGGCCTTCCGCCATCGCCGCGGCGCATCACCAGC 521
          | ||| || ||||| |  ||||||  ||| |||| || ||||| || ||||||||||
Db    427 GATTACAGTTATTTCTTTATCAATGGCGTTTGCTATAGCAGCCCGCTGGTATCACCAGC 486
```

Wed Mar 10 09:07:56 2004   us-10-033-109-3.rst

ORIGIN

```
Query Match         25.7%;   Score 437.6;  DB 14;  Length 984;
Best Local Similarity   69.1%;   Pred. No. 7e-39;
Matches 644;  Conservative   0;  Mismatches 279;  Indels   9;  Gaps   3;

Qy    88 CTTGTCCCGCGAACAACTGCCCAACTTCTCGGCCCAAACACCAGACGGCTCCGCCG 147
         ||||||||    |||   |||| |   ||||||    |||  ||||| ||||||||
Db    53 CTTGCTCAGCCGTCGGATCTCCTCCCACACTCGGC---GGCTCAGCCAATGCCACCGCGG 109

Qy   148 CCGCCTCCTTATCTGCGGCCATTCGCCCGCGTGGACAGCAAGTTCGTCGACACGGCCT 207
         |||||||||||||||||||||||||||||    |||  |||||||||||||||||
Db   110 CTGCAGAATTCATCTGCGGTGCTTCACAGCAATCTGGAATACCTCACAAACACCACCT 169

Qy   208 TCGGCCGTCGACAACACCTACCTCCTCTTTTCCGCCTACCTCGTTTTTCTATGCAGCTCG 267
         |||||||||||    |||||| |||| || ||||||||||||||||   |||||||||
Db   170 ACGCAGTGGACAACAACGTACCTCCTCTTCTCTGCTTCACCTGGTGTTTGCCATGCAGCTTG 229

Qy   268 GCTTCGCCATGCTCTCGCGCCGGCTCCGTCCGCGCCAAGAACACCATGAAACATCATGTCA 327
         |||||||||||||||||||||||| ||||||||||||||   ||||   ||| |||||
Db   230 GCTTTGCCATGTTGTGCGCGGGGCTCTGTCCGCGCCAAAAAAACACCATGAATATAATGCTTA 289

Qy   328 CCAACGTCCTGGACGCTGCCGCCGGCGGCTCCTTCTACTACCTCTTCGGCTTCGCCTTCG 387
         ||||||| |  |||   |||||  ||||||||| |||| |||||||||||||||||
Db   290 CGAACGTGCTTGATGCTGCACTGGTGGCCTTGTCTTATTACCTTTTTGGTTTTGCCTTTG 349

Qy   388 CTTTCGGCTTCCCCCTCCAAGGGCTTCATCGGTAAACATTTCTTCGGCCTCAAGGACAATCC 447
         ||||||     ||| |   |||| ||||| ||||   ||| |||||||  ||||||
Db   350 CCTTTGGAACTCCTTCCAATGGTTTCATTGGGAAACATTTCTTTGGGTTGAAAGAGTTTC 409

Qy   448 CTTCATCTCCTACGACTACAGCTACTTCCTCTCCACCAATGGGCCTTCGCCATGCCCGCCG 507
         |||| |||||||||||||||||||||||  ||||||| ||||  |||| |||||||||||
Db   410 CTTTACCATATTTGACTACAGCTACTTCTTATCAATGGGCTTTCGCCATAGCTGCTG 469

Qy   508 CCGGCATCACCAGCGGAAGCATCGCCGAACGCACACAGTTCGTGGCCTATCTCATCTACT 567
         ||||||||||||||||||| ||||  ||||||||||| ||||||||||||  || ||||
Db   470 CAGGTATCACCAGTGGGTCCATAGCTGAAAGGACACAATTTGTGGCGTACCTTATTTACT 529

Qy   568 CCTCCTTCCTCACCGGCTTCGTCTATCCGGTGGTCTCATCCGGTGGTCTCCCCAGACG 627
         ||||||   |||   |||| |  |||  || |||| ||||| ||||| ||||||||||
Db   530 CTTCCTTTTGACCGGTTTGTGTACCCTATTGTTTCACATGGTTCTCGGTCCGGTGACG 589

Qy   628 GCTGGGCCTCTGCCTTTAAGATCACCG---ACCGGCTATTTCCACCGGCGTAATAGACT 684
         ||||||||||      || || |       |  |||| ||||  ||||||| ||||||
Db   590 GTTGGGCCAGTGCATCAAAAACCGATGGAAACTTATTATTCGGTTCAGGAGTCATTGATT 649

Qy   685 TCGCCGGTTCCGGCGTAGTCCACATGGTCGCGCGGAATAGCCGGCCTATGGGGAGCGGCTGA 744
         ||||||||||   |||| |||||  |||| || ||||||||   ||||||||||||||||
Db   650 TTGCCGGTTCAGGTGTTGTTCATATAGTAGGAGGTATTGCCCGGTCTCATGGGGAGCTTTTA 709

Qy   745 TCGAAGGCCCAAGAATGGGACGTTTCGATCATGCAGGACGAGCGTGTGGCCTTGCGAGGGCC 804
         ||||||||      |||| |||| |||  |  ||||||   |  ||||||||  ||||||||
Db   710 TTGAAGGACCTAGAATCGGCCGGTTCGACCGGACCGGCTCTGTTGCATTAAGAGGTC 769

Qy   805 ACAGCGCGTCCTTAGTAGTCCTGGGAACCTTCTTGCTTGGTTCGGTTACGGGATTTA 864
         ||||||||||||| |||||||||||  ||| |||||||||||||||||||||||||||
Db   770 ATAGTGCATCACTAGTAGTATTGGGTACTTTCTTGCTTGGTTGGTTGGTATGGGTTTA 829

Qy   865 ACCCCGGTTCATTTAACAAAATCCTACTTACTTACCGGTAACTCA---GGAAATTACTACG 921
         ||||||||| ||||||||| |||  |||| || ||   |||| |     ||||| |||| |
Db   830 ACCCCGGTTCATTTTAACCATACTCCAAAAGTTATGATAGTCCAAATGGTGTTTATTATG 889

Qy   922 GTCAATGGAGCGCGGTTGGCAGAACCGCGGTCACCCTAGCCGGGTCAACAGCTG 981
         |||||||||||| ||||||||||||||||||  ||||||||||||||||||||||
Db   890 GTCAATGGAGTGCAATTGGAAGGACAGCTGTCACACGACATTGGCCGGTTGCACAGCTG 949

Qy   982 CCTTGACCACCGTATTCGGTAAACGGGTGATA 1013
         |  |||||||   ||||| ||| | |||
Db   950 CACTCCAACATTATTTAGCAAGAGACTTTTA 981
```

Search completed: March 9, 2004, 13:58:14
Job time : 4645 secs

```
                    GenCore version 5.1.6

              Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on:     March  9, 2004, 11:52:48 ; Search time 60 Seconds
                                (without alignments)
                         2354.563 Million cell updates/sec

Title:          US-10-033-109-4
Perfect score:  2639
Sequence:       1 MSLPACPAEQLAQLLGPNTT.........HGMQLRRVGPNASSTPTTDE 500

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters:    1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :     A_Geneseq_29Jan04:*
            1: geneseqp1980s:*
            2: geneseqp1990s:*
            3: geneseqp2000s:*
            4: geneseqp2001s:*
            5: geneseqp2002s:*
            6: geneseqp2003as:*
            7: geneseqp2003bs:*
            8: geneseqp2004s:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2639	100.0	500	6	AAO26533	Aao26533 Soybean h
2	2129.5	80.7	501	3	AAR74676	Aar74676 Arabidops
3	2129.5	80.7	501	3	AAG40620	Aag40620 Arabidops
4	2129.5	80.7	510	3	AAG36233	Aag36233 Arabidops
5	1954	74.0	494	6	AAO26534	Aao26534 Wheat hig
6	1943	73.6	446	4	AAG40621	Aag40621 Arabidops
7	1943	73.6	455	3	AAG36234	Aag36234 Arabidops
8	1913	72.5	440	4	AAG40622	Aag40622 Arabidops
9	1913	72.5	449	3	AAG36235	Aag36235 Arabidops
10	655.5	24.8	651	4	ABB62422	Abb62422 Drosophil
11	613	23.2	183	6	AAO26532	Aao26532 Corn high
12	491	18.6	401	5	ABB49278	Abb49278 Listeria
13	488	18.5	470	6	ADA34110	Ada34110 Acinetoba
14	481.5	18.2	431	5	ABP65520	Abp65520 Bifidobac
15	452.5	17.1	413	5	ABB54938	Abb54938 Lactococc
16	441	16.7	431	6	ADA32930	Ada32930 Acinetoba
17	430.5	16.3	438	4	AAG93227	Aag93227 C glutami
18	428	16.2	428	4	ABM70704	Abm70704 Photorhab
19	406.5	15.4	437	5	ABP39062	Abp39062 Staphyloc
20	403.5	15.3	488	6	ABP77916	Abp77916 N. gonorr
21	400	15.2	428	6	ADA89710	Ada89710 Staphyloc
22	400	15.2	428	6	ABM70882	Abm70882 Staphyloc
23	366	13.9	475	5	ABB92034	Abb92034 Herbicida
24	363.5	13.8	452	4	AAB76708	Aab76708 Corynebac
25	363.5	13.8	452	4	AAG93214	Aag93214 C glutami
26	335.5	12.7	470	6	AAO26538	Aao26538 Wheat hig
27	327.5	12.4	486	6	AAO26537	Aao26537 Soybean h
28	314	11.9	497	6	AAO26536	Aao26536 Rice high
29	301	11.4	924	4	ABG26085	Abg26085 Novel hum
30	204.5	7.7	339	5	ABP30411	Abp30411 Streptoco
31	204.5	7.7	347	5	ABP28641	Abp28641 Streptoco
32	202.5	7.7	347	5	ABP29900	Abp29900 Streptoco
33	193	7.3	223	7	ADC94047	Adc94047 E. faeciu
34	182	6.9	224	7	ADC94475	Adc94475 E. faeciu
35	173.5	6.6	441	5	ABP69263	Abp69263 Human pol
36	173.5	6.6	458	5	AAU78091	Aau78091 Human non
37	152	5.8	449	4	ABB57883	Abb57883 Drosophil
38	151	5.7	498	5	AAU78998	Aau78998 Mouse Rh
39	148	5.6	455	5	AAU78092	Aau78092 Mouse non
40	139	5.3	473	3	AAB42135	Aab42135 Human ORF
41	139	5.3	479	5	AAB29656	Aab29656 Human mem
42	139	5.3	479	4	AAB97000	Aab97000 Oesophagu
43	139	5.3	479	5	AAU78997	Aau78997 Human Rh
44	139	5.3	479	6	ABU56530	Abu56530 Lung canc
45	137	5.2	443	6	ABU31694	Abu31694 Protein e

```
                                     ALIGNMENTS

RESULT 1
AAO26533
ID   AAO26533 standard; protein; 500 AA.
XX
AC   AAO26533;
XX
DT   18-FEB-2003  (first entry)
XX
DE   Soybean high affinity ammonium transporter protein.
XX
KW   Herbicide; ammonium transporter protein; herbicide; transgenic plant;
KW   soybean.
XX
OS   Glycine max.
XX
XX   US2002142390-A1.
PN   US2002142390-A1.
XX
PD   03-OCT-2002.
XX
PF   28-DEC-2001; 2001US-00033109.
XX
PR   28-AUG-1998;    98US-0098248P.
PR   27-AUG-1999;    99US-00384625.
XX
PA   (ALLE/) ALLEN S M.
PA   (RAFA/) RAFALSKI J A.
XX
PI   Allen SM,   Rafalski JA;
XX
DR   WPI; 2003-102520/09.
DR   N-PSDB; AAL53985.
XX
PT   Novel ammonium transporter polypeptide useful for identifying enzymatic
PT   inhibitors, which is homologous to corn, soybean, wheat or rice ammonium
PT   transporter polypeptides.
XX
PS   Claim 14; Page 14-16; 27pp; English.
XX
CC   The invention relates to a novel ammonium transporter protein comprising
CC   90% homology based on the Clustal method compared to: a corn ammonium
CC   transporter polypeptide of 183 or 63 amino acids; a soybean ammonium
CC   transporter polypeptide of 500 or 486 amino acids; a wheat ammonium
CC   transporter polypeptide of 494 or 470 amino acids; and a rice ammonium
CC   transporter 497 amino acids fully defined in the specification. The
CC   isolated polynucleotide is useful for selecting an isolated
CC   polynucleotide that affects the level of expression of the ammonium
CC   transporter polypeptide in a plant cell. The ammonium transporter protein
CC   is useful for preparing antibodies which are useful for detecting the
```

```
CC   transporter protein in situ in cells or in vitro in cell extracts, and as
CC   targets to facilitate design and/or identify inhibitors of the enzymes
CC   that are useful as herbicides. The isolated polynucleotide is also useful
CC   for creating transgenic plants in which the polynucleotide is present at
CC   higher or lower levels than normal, and for designing and producing
CC   primer pairs which are useful in amplification or primer extension
CC   reactions. This sequence represents the soybean high affinity ammonium
CC   transporter protein of the invention
XX
SQ   Sequence 500 AA;

Query Match             100.0%;      Score 2639; DB 6;  Length 500;
Best Local Similarity   100.0%;      Pred. No. 2.9e-252;
Matches  500;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy    1 MSLPACPAEQLAQLLGPNTTDASAAASLICGHFAAVDSKFVDTAFAVDNTYLLFSAYLVF 60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MSLPACPAEQLAQLLGPNTTDASAAASLICGHFAAVDSKFVDTAFAVDNTYLLFSAYLVF 60

Qy   61 SMQLGFAMLCAGSVRAKNTMNIMLTNVLDAAAGGLFYLFGFAFAFGSPSNGFIGKHFFG 120
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 SMQLGFAMLCAGSVRAKNTMNIMLTNVLDAAAGGLFYLFGFAFAFGSPSNGFIGKHFFG 120

Qy  121 LKDIPSSSYDYSYFLYQWAFAIAAAGITSGSIAERTQFVAYLIYSSFLTGFVYPVVSHWF 180
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 LKDIPSSSYDYSYFLYQWAFAIAAAGITSGSIAERTQFVAYLIYSSFLTGFVYPVVSHWF 180

Qy  181 WSPDGWASAFKITDRLFSTGVIDFAGSGVVHMVGGIAGLWGALIEGPRMGRFDHAGRAVA 240
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 WSPDGWASAFKITDRLFSTGVIDFAGSGVVHMVGGIAGLWGALIEGPRMGRFDHAGRAVA 240

Qy  241 LRGHSASLVVLGTFLLMFGWYGFNPGSFNKILLTYGNSGNYYGQWSAVGRTAVTTTLAGS 300
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 LRGHSASLVVLGTFLLMFGWYGFNPGSFNKILLTYGNSGNYYGQWSAVGRTAVTTTLAGS 300

Qy  301 TAALTLFGKRVISGHWNVTDVCNGLLGGFAAITAGCSVVEPWAAIVCGFVASIVLIACN 360
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 TAALTLFGKRVISGHWNVTDVCNGLLGGFAAITAGCSVVEPWAAIVCGFVASIVLIACN 360

Qy  361 KLAEKVKFDDPLEAAQLHGGCGTWGVIFTALFAKKEYVGLGRAHGLLMGGGGKLLA 420
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 KLAEKVKFDDPLEAAQLHGGCGTWGVIFTALFAKKEYVGLGRAHGLLMGGGGKLLA 420

Qy  421 AHVIQILVIAGWVSATMGPLFWGLNKLKLLRISSEDELAGMDMTRHGGFAYAYEDDETHK 480
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 AHVIQILVIAGWVSATMGPLFWGLNKLKLLRISSEDELAGMDMTRHGGFAYAYEDDETHK 480

Qy  481 HGMQLRRVGPNASSTPTTDE 500
        ||||||||||||||||||||
Db  481 HGMQLRRVGPNASSTPTTDE 500

RESULT 2
AAR74676
ID   AAR74676 standard; protein; 501 AA.
XX
AC   AAR74676;
XX
DT   05-JAN-1996  (first entry)
XX
DE   Arabidopsis thaliana ammonium transporter.
XX
KW   Nitrogen metabolism; transgenic plant; ammonium transporter gene;
KW   antisense; inhibition; low input crop management.
XX
OS   Arabidopsis thaliana.
XX
PN   DE4337597-A1.
XX
PD   04-MAY-1995.
XX
PF   28-OCT-1993;    93DE-04337597.
XX
```

```
PR   28-OCT-1993;    93DE-04337597.
XX
PA   (GENB-) INST GENBIOLOGISCHE FORSCHUNG.
XX
PI   Frommer W, Ninnemann O;
XX
DR   WPI; 1995-171399/23.
DR   N-PSDB; AAQ88327.
XX
PT   Plant ammonium transporter DNA in sense or anti-sense orientation -
PT   useful for transforming cells, esp. plant cells for prodn. of transgenic
PT   plants with altered nitrogen metabolism.
XX
PS   Claim 3; Page 17-20; 23pp; German.
XX
CC   The ammonium transporter (AT) gene isolated from Arabidopsis thaliana
CC   (AAQ88327) is a preferred AT gene for use in generating transgenic plants
CC   ; the gene encodes the amino acid sequence in AAR74676. AT gene sequences
CC   can be incorporated into expression constructs in the sense orientation
CC   (for expression of translatable mRNA and hence synthesis of AT in the
CC   transgenic plants) or in the antisense orientation (resulting in
CC   antisense transcripts which inhibit synthesis of endogenous AT in the
CC   transgenic plants). Nitrogen metabolism is altered in the resulting
CC   transgenic plants; the changes in AT activity may also result in plants
CC   which are suitable for "low input" crop management or which can be grown
CC   in acid soils
XX
SQ   Sequence 501 AA;

Query Match            80.7%;      Score 2129.5;  DB 2;   Length 501;
Best Local Similarity  80.9%;      Pred. No. 8.6e-202;
Matches  407;  Conservative   42;  Mismatches   39;  Indels   15;  Gaps   7;

Qy    5 ACPAEQLAQLLGPNTTDASAAASLICGHFAAVDSKFVDTAFAVDNTYLLFSAYLVFSMQL 64
        :|:||||||||:|||  |:|:::||||||||:|:||||||||||||||:||||||||||||
Db    2 SCSATDLAVLLGPN--ATAAANYICGQLGDVNNKFIDTAFAIDNTYLLFSAYLVFSMQL 58

Qy   65 GFAMLCAGSVRAKNTMNIMLTNVLDAAAGGLFYLFGFAPAFGSPSNGFIGKHFFGLKDI 124
        ||||||||||||||||||||||||||||||||||||||:||||||||||:|:||||||
Db   59 GFAMLCAGSVRAKNTMNIMLTNVLDAAAGGLFYLFGYAPAFGSPSNGFIGKHYFGLKDI 118

Qy  125 PSSSYDYSYFLYQWAFAIAAAGITSGSIAERTQFVAYLIYSSFLTGFVYPVVSHWFWSPD 184
        |:::||:|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  119 PTASADYSNFLYQWAFAIAAAGITSGSIAERTQFVAYLIYSSFLTGFVYPVVSHWFWSVD 178

Qy  185 GWASAFKIT-DRLFSTGVIDFAGSGVVHMVGGIAGLWGALIEGPRMGRFDHAGRAVALRG 243
        ||||:|:|| |||||||:|||||||||||||||||||||||||||:|||:|||||:|:||
Db  179 GWASPFRTDGDLLFSTGAIDFAGSGVVHMVGGIAGLWGALIEGPRLGRFDNGGRAIALRG 238

Qy  244 HSASLVVLGTFLLMFGWYGFNPGSFNKILLTYGNSGNYYGQWSAVGRTAVTTTLAGSTAA 303
        |||||||||||||||||||||||||||||:||| :||.:|||||||||||:|:|||:|
Db  239 HSASLVVLGTFLLMFGWYGFNPGSFNKILVTY-BTGTYNGQWSAVGRTAVTTTLAGCTAA 297

Qy  304 LTTLFGKRVISGHWNVTDVCNGLLGGFAAITAGCSVVEPWAAIVCGFVASIVLIACNKLA 363
        ||||||||::|||||||||||||||||||||:|||||||||||::|:|::|:||:|||||
Db  298 LTTLFGKRLLSGHWNVTDVCNGLLGGFAAITGGCSVVEPWAAIICGFVAALVLLGCNKLA 357

Qy  364 EKVKFDDPLEAAQLHGGCGTWGVIFTALFAKKEYVKEVVG--LGRAHGLLMGGGGKLLAA 421
        ||:|:|||||||||||||||:||:||||||::|:  |:: ||:||:||||:||||||:|
Db  358 EKLKYDDPLEAAQLHGGCGAWGLIFTALFAQEKYLNQIYGNKPGRPHGLFMGGGGKLLGA 417

Qy  422 HVIQILVIAGWVSATMGPLFWGLNKLKLLRISSEDELAGMDMTRHGGFAYAY-EDDETHK 480
        :|||||||:|:||:||||| ::: ||||||||||||:||||||||||:|:::|:: |:|
Db  418 QLIQIVITGWVSATMGTLFFIKMKLLRISSEDEMAGMDMTRHGGFAYMYFDDDESHK 477

Qy  481 HGMQLRRVGP-----NASSTPT 497
        |||||||| :     |||:|||
Db  478 -AIQLRRVEPRSPSPGGANTTPT 499

RESULT 3
AAG40620
ID   AAG40620 standard; protein; 501 AA.
```

XX
AC   AAG40620;
XX
DT   18-OCT-2000 (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 50426.
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-00301439.
XX

PR	25-FEB-1999;	99US-0121825P.
PR	05-MAR-1999;	99US-0123180P.
PR	09-MAR-1999;	99US-0123548P.
PR	23-MAR-1999;	99US-0125788P.
PR	25-MAR-1999;	99US-0126264P.
PR	29-MAR-1999;	99US-0126785P.
PR	01-APR-1999;	99US-0127462P.
PR	06-APR-1999;	99US-0128234P.
PR	08-APR-1999;	99US-0128714P.
PR	16-APR-1999;	99US-0129845P.
PR	19-APR-1999;	99US-0130077P.
PR	21-APR-1999;	99US-0130449P.
PR	23-APR-1999;	99US-0130510P.
PR	23-APR-1999;	99US-0130891P.
PR	28-APR-1999;	99US-0131449P.
PR	30-APR-1999;	99US-0132048P.
PR	30-APR-1999;	99US-0132407P.
PR	04-MAY-1999;	99US-0132484P.
PR	05-MAY-1999;	99US-0132485P.
PR	06-MAY-1999;	99US-0132486P.
PR	06-MAY-1999;	99US-0132487P.
PR	07-MAY-1999;	99US-0132663P.
PR	11-MAY-1999;	99US-0134256P.
PR	14-MAY-1999;	99US-0134218P.
PR	14-MAY-1999;	99US-0134219P.
PR	14-MAY-1999;	99US-0134221P.
PR	14-MAY-1999;	99US-0134370P.
PR	18-MAY-1999;	99US-0134768P.
PR	19-MAY-1999;	99US-0134941P.
PR	20-MAY-1999;	99US-0135124P.
PR	21-MAY-1999;	99US-0135353P.
PR	24-MAY-1999;	99US-0135629P.
PR	25-MAY-1999;	99US-0136021P.
PR	27-MAY-1999;	99US-0136392P.
PR	28-MAY-1999;	99US-0136782P.
PR	01-JUN-1999;	99US-0137222P.
PR	03-JUN-1999;	99US-0137528P.
PR	04-JUN-1999;	99US-0137502P.
PR	07-JUN-1999;	99US-0137724P.
PR	08-JUN-1999;	99US-0138094P.
PR	10-JUN-1999;	99US-0138540P.
PR	10-JUN-1999;	99US-0138847P.
PR	14-JUN-1999;	99US-0139119P.
PR	16-JUN-1999;	99US-0139452P.
PR	16-JUN-1999;	99US-0139453P.
PR	17-JUN-1999;	99US-0139492P.
PR	18-JUN-1999;	99US-0139454P.
PR	18-JUN-1999;	99US-0139455P.
PR	18-JUN-1999;	99US-0139456P.
PR	18-JUN-1999;	99US-0139457P.
PR	18-JUN-1999;	99US-0139458P.
PR	18-JUN-1999;	99US-0139459P.
PR	18-JUN-1999;	99US-0139460P.
PR	18-JUN-1999;	99US-0139461P.
PR	18-JUN-1999;	99US-0139462P.
PR	18-JUN-1999;	99US-0139463P.
PR	18-JUN-1999;	99US-0139750P.
PR	18-JUN-1999;	99US-0139763P.
PR	21-JUN-1999;	99US-0139817P.
PR	22-JUN-1999;	99US-0139899P.
PR	23-JUN-1999;	99US-0140353P.
PR	23-JUN-1999;	99US-0140354P.
PR	24-JUN-1999;	99US-0140695P.
PR	28-JUN-1999;	99US-0140823P.
PR	29-JUN-1999;	99US-0140991P.
PR	30-JUN-1999;	99US-0141287P.
PR	01-JUL-1999;	99US-0141842P.
PR	01-JUL-1999;	99US-0142154P.
PR	02-JUL-1999;	99US-0142055P.
PR	06-JUL-1999;	99US-0142390P.
PR	08-JUL-1999;	99US-0142803P.
PR	09-JUL-1999;	99US-0142920P.
PR	12-JUL-1999;	99US-0142977P.
PR	13-JUL-1999;	99US-0143542P.
PR	14-JUL-1999;	99US-0143624P.
PR	15-JUL-1999;	99US-0144005P.
PR	16-JUL-1999;	99US-0140085P.
PR	16-JUL-1999;	99US-0144086P.
PR	19-JUL-1999;	99US-0144325P.
PR	19-JUL-1999;	99US-0144331P.
PR	19-JUL-1999;	99US-0144332P.
PR	19-JUL-1999;	99US-0144333P.
PR	19-JUL-1999;	99US-0144334P.
PR	19-JUL-1999;	99US-0144335P.
PR	20-JUL-1999;	99US-0144352P.
PR	20-JUL-1999;	99US-0144632P.
PR	21-JUL-1999;	99US-0144884P.
PR	21-JUL-1999;	99US-0144814P.
PR	21-JUL-1999;	99US-0145086P.
PR	22-JUL-1999;	99US-0150088P.
PR	22-JUL-1999;	99US-0150085P.
PR	22-JUL-1999;	99US-0150087P.
PR	22-JUL-1999;	99US-0150089P.
PR	22-JUL-1999;	99US-0145192P.
PR	23-JUL-1999;	99US-0145145P.
PR	23-JUL-1999;	99US-0145218P.
PR	26-JUL-1999;	99US-0145224P.
PR	27-JUL-1999;	99US-0145276P.
PR	27-JUL-1999;	99US-0145913P.
PR	27-JUL-1999;	99US-0145918P.
PR	27-JUL-1999;	99US-0145919P.
PR	28-JUL-1999;	99US-0145951P.
PR	02-AUG-1999;	99US-0146386P.
PR	02-AUG-1999;	99US-0146388P.
PR	03-AUG-1999;	99US-0146389P.
PR	03-AUG-1999;	99US-0147038P.
PR	04-AUG-1999;	99US-0147204P.
PR	04-AUG-1999;	99US-0147302P.
PR	05-AUG-1999;	99US-0147192P.
PR	05-AUG-1999;	99US-0147260P.
PR	06-AUG-1999;	99US-0147303P.
PR	06-AUG-1999;	99US-0147416P.
PR	09-AUG-1999;	99US-0147493P.
PR	09-AUG-1999;	99US-0147935P.
PR	10-AUG-1999;	99US-0149175P.
PR	11-AUG-1999;	99US-0148171P.
PR	12-AUG-1999;	99US-0148319P.
PR	13-AUG-1999;	99US-0148341P.
PR	13-AUG-1999;	99US-0148565P.
PR	16-AUG-1999;	99US-0148684P.
PR	17-AUG-1999;	99US-0149368P.
PR	18-AUG-1999;	99US-0149175P.
PR	20-AUG-1999;	99US-0149426P.
PR	20-AUG-1999;	99US-0149722P.
PR	20-AUG-1999;	99US-0149723P.
PR	23-AUG-1999;	99US-0149929P.
PR	23-AUG-1999;	99US-0149902P.
PR	23-AUG-1999;	99US-0149930P.

```
PR    25-AUG-1999;    99US-0150566P.
PR    26-AUG-1999;    99US-0150884P.
PR    27-AUG-1999;    99US-0151065P.
PR    27-AUG-1999;    99US-0151066P.
PR    27-AUG-1999;    99US-0151080P.
PR    30-AUG-1999;    99US-0151303P.
PR    31-AUG-1999;    99US-0151433P.
PR    01-SEP-1999;    99US-0151930P.
PR    07-SEP-1999;    99US-0152363P.
PR    10-SEP-1999;    99US-0153070P.
PR    13-SEP-1999;    99US-0153758P.
PR    15-SEP-1999;    99US-0154018P.
PR    16-SEP-1999;    99US-0154039P.
PR    20-SEP-1999;    99US-0154779P.
PR    22-SEP-1999;    99US-0155139P.
PR    23-SEP-1999;    99US-0155486P.
PR    24-SEP-1999;    99US-0155659P.
PR    28-SEP-1999;    99US-0156458P.
PR    29-SEP-1999;    99US-0156596P.
PR    04-OCT-1999;    99US-0157117P.
PR    05-OCT-1999;    99US-0157753P.
PR    06-OCT-1999;    99US-0157865P.
PR    07-OCT-1999;    99US-0158029P.
PR    08-OCT-1999;    99US-0158232P.
PR    12-OCT-1999;    99US-0158369P.
PR    13-OCT-1999;    99US-0159293P.
PR    13-OCT-1999;    99US-0159294P.
PR    13-OCT-1999;    99US-0159295P.
PR    14-OCT-1999;    99US-0159329P.
PR    14-OCT-1999;    99US-0159330P.
PR    14-OCT-1999;    99US-0159331P.
PR    14-OCT-1999;    99US-0159637P.
PR    14-OCT-1999;    99US-0159638P.
PR    18-OCT-1999;    99US-0159584P.
PR    21-OCT-1999;    99US-0160741P.
PR    21-OCT-1999;    99US-0160767P.
PR    21-OCT-1999;    99US-0160768P.
PR    21-OCT-1999;    99US-0160770P.
PR    21-OCT-1999;    99US-0160814P.
PR    21-OCT-1999;    99US-0160815P.
PR    22-OCT-1999;    99US-0160980P.
PR    22-OCT-1999;    99US-0160981P.
PR    22-OCT-1999;    99US-0160989P.
PR    25-OCT-1999;    99US-0161404P.
PR    25-OCT-1999;    99US-0161405P.
PR    25-OCT-1999;    99US-0161406P.
PR    25-OCT-1999;    99US-0161359P.
PR    26-OCT-1999;    99US-0161360P.
PR    26-OCT-1999;    99US-0161361P.
PR    28-OCT-1999;    99US-0161920P.
PR    28-OCT-1999;    99US-0161992P.
PR    28-OCT-1999;    99US-0161993P.
PR    29-OCT-1999;    99US-0162142P.
```

```
Qy   244  HSASLVLVGTFLLWFGWYGFNPGSFNKILLTYGNSGNYYGQWSAVGRTAVTTLAGSTAA  303
          |||||||||||||||||||||||||||||    ::|||:||||||||||||||||||||
Db   239  HSASLVVLGTFLLWFGWYGFNPGSFNKILLVTY-ETGTYNGQWSAVGRTAVTTLAGCTAA  297

Qy   304  LTTLFGKRVISGHWNVTDVCNGLLGGFAAITAGCSVVEPWAAIVCGFVASIVLIACNKLA  363
          |||||||  :|||||||||||||||||||||||||||||||||:: |||: ||| |||||
Db   298  LTTLFGKRLLSGHWNVTDVCNGLLGGFAAITGGCSVVEPWAAIICGFVAALVLLGCNKLA  357

Qy   364  EKVKFDDPLEAAQLHGGCGTWGVIFTALFAKKEYVKEVYG--LGRAHGLLMCGGGKLLAA  421
          :|:|||||||||||||||||  ||||||||||||:: ::||  || |||:||||:||||
Db   358  EKLKYDDPLEAAQLHGGCGAWGLIFTALFAQEKYLNQIYGNKPGRPHGLFMGGGGKLLGA  417

Qy   422  HVIQILVIAGWVSATMGPLFWGLNKLKLLRISSEDELAGMDTRHGGFAYAY-EDDETHK   480
          ||:|||||||||||||||||||||:|||||||||||||||||||||||: | : |||||
Db   418  QLIQIVITGWVSATMGTLFFILKKMKLLRISSEDEMAGMDTRHGGFAYMYFDDDESHK   477

Qy   481  HGMQLRRVGP------NASSTPT  497
               |||||     ::|||
Db   478  -AIQLRRVEPRSPSGANTTPT   499
```

```
RESULT 4
AAG36233
ID   AAG36233 standard; protein; 510 AA.
XX
AC   AAG36233;
XX
DT   18-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 44375.
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000;  2000EP-00301439.
XX
PR   25-FEB-1999;    99US-0121825P.
PR   05-MAR-1999;    99US-0123180P.
PR   09-MAR-1999;    99US-0123548P.
PR   23-MAR-1999;    99US-0125788P.
PR   25-MAR-1999;    99US-0126264P.
PR   29-MAR-1999;    99US-0127462P.
PR   01-APR-1999;    99US-0127462P.
PR   06-APR-1999;    99US-0128234P.
PR   08-APR-1999;    99US-0128714P.
PR   16-APR-1999;    99US-0129845P.
PR   19-APR-1999;    99US-0130077P.
PR   21-APR-1999;    99US-0130449P.
PR   23-APR-1999;    99US-0130510P.
PR   23-APR-1999;    99US-0130891P.
PR   28-APR-1999;    99US-0131449P.
PR   30-APR-1999;    99US-0132048P.
PR   30-APR-1999;    99US-0132407P.
PR   04-MAY-1999;    99US-0132484P.
PR   05-MAY-1999;    99US-0132485P.
PR   06-MAY-1999;    99US-0132486P.
PR   06-MAY-1999;    99US-0132487P.
PR   07-MAY-1999;    99US-0132863P.
PR   11-MAY-1999;    99US-0134256P.
PR   14-MAY-1999;    99US-0134218P.
PR   14-MAY-1999;    99US-0134219P.
PR   14-MAY-1999;    99US-0134221P.
PR   14-MAY-1999;    99US-0134370P.
PR   18-MAY-1999;    99US-0134768P.
PR   19-MAY-1999;    99US-0134941P.
```

```
Query Match         80.7%;   Score 2129.5;  DB 3;   Length 501;
Best Local Similarity   80.9%;   Pred. No. 8.6e-202;
Matches  407;  Conservative  42;  Mismatches  39;   Indels  15;   Gaps   7;

Qy     5  ACPAEQLAQLLGPNTTDASAAASLICGHFAAVDSKFVDTAFAVDNTYLLFSAYLVFSMQL  64
          :| ||||||  :||||    :||:|: | |||:::||||||||||||||||||||||||||
Db     2  SCSATDLAVLLGPN---ATAAANYICGQLGDVNNKPIDTAFAIDNTYLLFSAYLVFSMQL  58

Qy    65  GFAMLCAGSVRAKNTMNIMLTNVLDAAAGGLFYYLPGFAFAFGSPSNGFIGKHFFGLKDI  124
          |||||||||||||||||||||||||||||||||||:||||||||||||||||:|||||||
Db    59  GFAMLCAGSVRAKNTMNIMLTNVLDAAAGGLFYYLPGYAFAFGSPSNGFIGKHYFGLKDI  118

Qy   125  PSSSYDYSFLYQWAFAIAAAGITSGSIAERTQFVAYLIYSSFLTGFVYPVVSHWFWSPD  184
          |::::||  ||||| |||||||||||||||||||||||||||||||||||||||||:|||
Db   119  PTASADYSNFLYQWAFAIAAAGITSGSIAERTQFVAYLIYSSFLTGFVYPVVSHWFWSVD  178

Qy   185  GWASAFKIT-DRLFSTGVIDFAGSGVVHMVGGIAGLWGALIEGPRMGRFDHAGRAVALRG  243
          ||||:|   | || |||:||||||||||||||||||||||||||||: |||| |||:|||
Db   179  GWASPFKTDGDLLFSTGAIDFAGSGVVHMVGGIAGLWGALIEGPRLGRFDNGGRAIALRG  238
```

PR	20-MAY-1999;	99US-0135124P.	PR	02-AUG-1999;	99US-0146386P.
PR	21-MAY-1999;	99US-0135353P.	PR	02-AUG-1999;	99US-0146388P.
PR	24-MAY-1999;	99US-0135629P.	PR	02-AUG-1999;	99US-0146389P.
PR	25-MAY-1999;	99US-0136021P.	PR	03-AUG-1999;	99US-0147038P.
PR	27-MAY-1999;	99US-0136392P.	PR	04-AUG-1999;	99US-0147204P.
PR	28-MAY-1999;	99US-0137782P.	PR	04-AUG-1999;	99US-0147302P.
PR	01-JUN-1999;	99US-0137222P.	PR	05-AUG-1999;	99US-0147192P.
PR	03-JUN-1999;	99US-0137528P.	PR	05-AUG-1999;	99US-0147260P.
PR	04-JUN-1999;	99US-0137502P.	PR	06-AUG-1999;	99US-0147303P.
PR	07-JUN-1999;	99US-0137724P.	PR	06-AUG-1999;	99US-0147416P.
PR	08-JUN-1999;	99US-0138094P.	PR	09-AUG-1999;	99US-0147493P.
PR	10-JUN-1999;	99US-0138540P.	PR	09-AUG-1999;	99US-0147935P.
PR	10-JUN-1999;	99US-0138847P.	PR	10-AUG-1999;	99US-0148171P.
PR	14-JUN-1999;	99US-0139119P.	PR	11-AUG-1999;	99US-0148819P.
PR	16-JUN-1999;	99US-0139452P.	PR	12-AUG-1999;	99US-0148341P.
PR	16-JUN-1999;	99US-0139453P.	PR	13-AUG-1999;	99US-0148565P.
PR	17-JUN-1999;	99US-0139492P.	PR	13-AUG-1999;	99US-0148684P.
PR	18-JUN-1999;	99US-0139454P.	PR	16-AUG-1999;	99US-0149368P.
PR	18-JUN-1999;	99US-0139455P.	PR	17-AUG-1999;	99US-0149175P.
PR	18-JUN-1999;	99US-0139456P.	PR	18-AUG-1999;	99US-0149426P.
PR	18-JUN-1999;	99US-0139457P.	PR	20-AUG-1999;	99US-0149722P.
PR	18-JUN-1999;	99US-0139458P.	PR	20-AUG-1999;	99US-0149723P.
PR	18-JUN-1999;	99US-0139459P.	PR	20-AUG-1999;	99US-0149929P.
PR	18-JUN-1999;	99US-0139460P.	PR	23-AUG-1999;	99US-0149902P.
PR	18-JUN-1999;	99US-0139461P.	PR	23-AUG-1999;	99US-0149930P.
PR	18-JUN-1999;	99US-0139462P.	PR	25-AUG-1999;	99US-0150566P.
PR	18-JUN-1999;	99US-0139463P.	PR	26-AUG-1999;	99US-0150884P.
PR	18-JUN-1999;	99US-0139750P.	PR	27-AUG-1999;	99US-0151065P.
PR	18-JUN-1999;	99US-0139763P.	PR	27-AUG-1999;	99US-0151066P.
PR	21-JUN-1999;	99US-0139817P.	PR	30-AUG-1999;	99US-0151080P.
PR	22-JUN-1999;	99US-0139899P.	PR	31-AUG-1999;	99US-0151303P.
PR	23-JUN-1999;	99US-0140353P.	PR	01-SEP-1999;	99US-0151438P.
PR	23-JUN-1999;	99US-0140354P.	PR	07-SEP-1999;	99US-0151930P.
PR	24-JUN-1999;	99US-0140695P.	PR	10-SEP-1999;	99US-0152363P.
PR	28-JUN-1999;	99US-0142390P.	PR	11-SEP-1999;	99US-0153070P.
PR	29-JUN-1999;	99US-0140991P.	PR	15-SEP-1999;	99US-0153758P.
PR	30-JUN-1999;	99US-0141287P.	PR	16-SEP-1999;	99US-0154018P.
PR	01-JUL-1999;	99US-0141842P.	PR	16-SEP-1999;	99US-0154039P.
PR	01-JUL-1999;	99US-0142154P.	PR	20-SEP-1999;	99US-0154779P.
PR	02-JUL-1999;	99US-0142055P.	PR	22-SEP-1999;	99US-0155139P.
PR	06-JUL-1999;	99US-0142390P.	PR	23-SEP-1999;	99US-0155486P.
PR	08-JUL-1999;	99US-0142803P.	PR	24-SEP-1999;	99US-0155659P.
PR	09-JUL-1999;	99US-0142920P.	PR	28-SEP-1999;	99US-0156458P.
PR	12-JUL-1999;	99US-0142977P.	PR	29-SEP-1999;	99US-0155596P.
PR	13-JUL-1999;	99US-0143542P.	PR	04-OCT-1999;	99US-0157117P.
PR	14-JUL-1999;	99US-0143624P.	PR	05-OCT-1999;	99US-0157753P.
PR	15-JUL-1999;	99US-0144333P.	PR	06-OCT-1999;	99US-0157865P.
PR	16-JUL-1999;	99US-0144005P.	PR	07-OCT-1999;	99US-0158029P.
PR	16-JUL-1999;	99US-0144085P.	PR	08-OCT-1999;	99US-0158232P.
PR	16-JUL-1999;	99US-0144086P.	PR	12-OCT-1999;	99US-0158369P.
PR	19-JUL-1999;	99US-0144325P.	PR	13-OCT-1999;	99US-0159293P.
PR	19-JUL-1999;	99US-0144331P.	PR	13-OCT-1999;	99US-0159294P.
PR	19-JUL-1999;	99US-0144332P.	PR	14-OCT-1999;	99US-0159295P.
PR	19-JUL-1999;	99US-0144333P.	PR	14-OCT-1999;	99US-0159329P.
PR	19-JUL-1999;	99US-0144334P.	PR	14-OCT-1999;	99US-0159330P.
PR	19-JUL-1999;	99US-0144335P.	PR	14-OCT-1999;	99US-0159331P.
PR	20-JUL-1999;	99US-0144352P.	PR	14-OCT-1999;	99US-0159637P.
PR	20-JUL-1999;	99US-0144632P.	PR	18-OCT-1999;	99US-0159638P.
PR	21-JUL-1999;	99US-0144884P.	PR	21-OCT-1999;	99US-0159584P.
PR	21-JUL-1999;	99US-0145192P.	PR	21-OCT-1999;	99US-0160741P.
PR	22-JUL-1999;	99US-0145145P.	PR	21-OCT-1999;	99US-0160767P.
PR	22-JUL-1999;	99US-0145218P.	PR	21-OCT-1999;	99US-0160768P.
PR	23-JUL-1999;	99US-0145224P.	PR	21-OCT-1999;	99US-0160770P.
PR	23-JUL-1999;	99US-0144814P.	PR	22-OCT-1999;	99US-0160814P.
PR	26-JUL-1999;	99US-0145086P.	PR	22-OCT-1999;	99US-0160815P.
PR	27-JUL-1999;	99US-0145088P.	PR	22-OCT-1999;	99US-0160980P.
PR	27-JUL-1999;	99US-0145913P.	PR	25-OCT-1999;	99US-0160981P.
PR	27-JUL-1999;	99US-0145918P.	PR	25-OCT-1999;	99US-0160989P.
PR	27-JUL-1999;	99US-0145919P.	PR	25-OCT-1999;	99US-0161404P.
PR	28-JUL-1999;	99US-0145951P.	PR	25-OCT-1999;	99US-0161405P.
			PR	25-OCT-1999;	99US-0161406P.
			PR	26-OCT-1999;	99US-0161359P.
			PR	26-OCT-1999;	99US-0161360P.

```
PR   26-OCT-1999;   99US-0161361P.
PR   28-OCT-1999;   99US-0161920P.
PR   28-OCT-1999;   99US-0161992P.
PR   29-OCT-1999;   99US-0161993P.
PR   29-OCT-1999;   99US-0162142P.

Query Match             80.7%;  Score 2129.5;  DB 3;  Length 510;
Best Local Similarity   80.9%;  Pred. No. 8.8e-202;
Matches  407;  Conservative  42;  Mismatches  39;  Indels  15;  Gaps  7;

Qy    5 ACPAEQLAQLLGPNTTDASAAASLICGHFAAVDSKFVDTAPAVDNTYLLFSAYLVFSMQL 64
        : || | |||||          :|||: :|||: ||||||||||||||||||||||||||
Db    2 SCSATDLAVLLGPN---ATAAANYICGQLGDVNNKFIDTAFAIDNTYLLFSAYLVFSMQL 58

Qy   65 GFAMLCAGSVRAKNTMNIMLTNVLDAAAGLFYYLFGFAFAFGSPSNGFIGKHFFGLKDI 124
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||
Db   59 GFAMLCAGSVRAKNTMNIMLTNVLDAAAGLFYYLFGFAFAFGSPSNGFIGKHYFGLKDI 118

Qy  125 PSSSYDYSYFLYQWAFAIAAAGITSGSIAERTQFVAYLIYSSFLTGFVYPVVSHWFWSPD 184
        |:|: |||| :||||||||||||||||||||||||||||||||||||||||||||||| |
Db  119 PTASADYSNFLYQWAFAIAAAGITSGSIAERTQFVAYLIYSSFLTGFVYPVVSHWFWSVD 178

Qy  185 GWASAFKIT-DRLFSTGVIDFAGSGVVHMVGGIAGLWGALIEGPRMGRFDHAGRAVALRG 243
        ||| :  | ::|||||.:|||||||||||||||||||||||||||:|||: :||::||||
Db  179 GWASPFRTDGDLLFSTGAIDFAGSGVVHMVGGIAGLWGALIEGPRLGRFDNGGRAIALRG 238

Qy  244 HSASLVVLGTFLLWFGWYGFNPGSFNKILLTYGNSGNYYGQWSAVGRTAVTTTLAGSTAA 303
        ||||||||||||||||||||||||||||| :| : | ||||||||||||||||||.|||
Db  239 HSASLVVLGTFLLWFGWYGFNPGSFNKILVTY-ETGTYNGQWSAVGRTAVTTTLAGCTAA 297

Qy  304 LTTLFGKRVISGHWNVTDVCNGLLGGFAAITAGCSVVEPWAAIVCGFVASIVLIACNKLA 363
        ||||||||::|||||||||||||||||||||.|||||||||||:||||:: :|:.||||
Db  298 LTTLFGKRLLSGHWNVTDVCNGLLGGFAAITGGCSVVEPWAAIICGFVAALVLLGCNKLA 357

Qy  364 EKVKFDDPLEAAQLHGGCGTWGVIFTALFAKKEYVKEVYG--LGRAHGLLMGGGGKLLAA 421
        ||:|:|||||||||||||||.||:|||||||: || ::||  :|| :||| |||||||.|
Db  358 EKLKYDDPLEAAQLHGGCGAWGLIFTALFAQEKYLNQIYGNKPGRPHGLFMGGGGKLLGA 417

Qy  422 HVIQILVIAGWVSATMGPLFWGLNKLKLLRISSEDELAGMDMTRHGGFAYAY-EDDETHK 480
        ::|||:||.|||||||| ||: |:| ||||||||||:|||||||||||||.| :||:: |
Db  418 QLIQIVITGWVSATMGTLFFILKKMKLLRISSEDEMAGMDMTRHGGFAYMYFDDDDESHK 477

Qy  481 HGMQLRRVGP------NASSTPT 497
        :.:||||| |      .||:|||
Db  478 -AIQLRRVEPRSPSPSGANTTPT 499
```

PT   Novel ammonium transporter polypeptide useful for identifying enzymatic
PT   inhibitors, which is homologous to corn, soybean, wheat or rice ammonium
PT   transporter polypeptides.
XX
PS   Claim 14; Page 17-18; 27pp; English.
XX
CC   The invention relates to a novel ammonium transporter protein comprising
CC   90% homology based on the Clustal method compared to: a corn ammonium
CC   transporter polypeptide of 183 or 63 amino acids; a soybean ammonium
CC   transporter polypeptide of 500 or 486 amino acids; a wheat ammonium
CC   transporter polypeptide of 494 or 470 amino acids; and a rice ammonium
CC   transporter 497 amino acids fully defined in the specification. The
CC   isolated polynucleotide is useful for selecting an isolated
CC   polynucleotide that affects the level of expression of the ammonium
CC   transporter polypeptide in a plant cell. The ammonium transporter protein
CC   is useful for preparing antibodies which are useful for detecting the
CC   transporter protein in situ in cells or in vitro in cell extracts, and as
CC   targets to facilitate design and/or identify inhibitors of the enzymes
CC   that are useful as herbicides. The isolated polynucleotide is also useful
CC   for creating transgenic plants in which the polynucleotide is present at
CC   higher or lower levels than normal, and for designing and producing
CC   primer pairs which are useful in amplification or primer extension
CC   reactions. This sequence represents the wheat high affinity ammonium
CC   transporter protein of the invention
XX
SQ   Sequence 494 AA;

```
Query Match             74.0%;  Score 1954;  DB 6;  Length 494;
Best Local Similarity   74.8%;  Pred. No. 2e-184;
Matches  374;  Conservative  46;  Mismatches  62;  Indels  18;  Gaps  5;

Qy    3 LPACPAEQLAQLLGPNTTDASAAASLICGHFAAVDSKFVDTAFAVDNTYLLFSAYLVFSM 62
         :| ||||:|||| |  : |:  |:||||:|:        :|| ||||||||||||||||
Db    1 MSATCAADLGPLLG---AAAANATDYICNRFA------DTTSAVDSTYLLFSAXLVFAM 50

Qy   63 QLGFAMLCAGSVRAKNTMNIMLTNVLDAAAGLFYYLFGFAFAFGSPSNGFIGKHFFGLK 122
        |||||||||||||||||||||||||||||||||||||||||||.|||||||||||||||
Db   51 QLGFAMLCAGSVRAKNTMNIMLTNVLDAAAGLFYYLFGFAFAFGTPSNGFIGKHFFGLK 110

Qy  123 DIPSSSYDYSYFLYQWAFAIAAAGITSGSIAERTQFVAYLIYSSFLTGFVYPVVSHWFWS 182
        |:|  :  ||:||.|||||||||||||||||||||||||||||:||||||||||||| |
Db  111 DMPQTGFDYSFFLFQWAFAIAAAGITSGSIAERTQFVAYLIYSAFLTGFVYPVVSHWIWS 170

Qy  183 PDGWASAFKIT-DRLFSTGVIDFAGSGVVHMVGGIAGLWGALIEGPRMGRFDHAGRAVAL 241
         ||||||  |: ||| |:||||||||||||||||||| ||||||||::||||||||:||
Db  171 VDGWASAARTSGPLLFKSGVIDFAGSGVVHMVGGIAGFWGALIEGPRIGRFDHAGRSVAL 230

Qy  242 RGHSASLVVLGTFLLWFGWYGFNPGSFNKILLTYGNSGNYYGQWSAVGRTAVTTTLAGST 301
        :|||||||||||||||||||||||||| :|: :||  ||:||||| |||||||||||||
Db  231 KGHSASLVVLGTFLLWFGWYGFNPGSFVTILKSYGPPGSINGQWSGVGRTAVTTTLAGSV 290

Qy  302 AALTTLFGKRVISGHWNVTDVCNGLLGGFAAITAGCSVVEPWAAIVCGFVASIVLIACNK 361
        ||||||||||::.|||||.||||||||||||||||||||:||||:||||::.:|||  |
Db  291 AALTTLFGKRLQTGHWNVVDVCNGLLGGFAAITAGCSVVDPWAAIVCGFVSAWVLIGLNA 350

Qy  362 LAEKVKFDDPLEAAQLHGGCGTWGVIFTALFAKKEYVKEVYGLGRAHGLLMGGGGKLLAA 421
        || :|:.||||||||||||||.||:|||||||:: ||.:||.||: ||| |||||:||||
Db  351 LAGRLKYDDPLEAAQLHGGCGAWGIIFTALFAKKQYVEIYGAGRPYGLLFLGGGGRLLAA 410

Qy  422 HVIQILVIAGWVSATMGPLFWGLNKLKLLRISSEDELAGMDMTRHGGFAYAYEDDETHKH 481
        |::||||||| || ||||||  |:| .|||||:|||:||||:||||||||.|.||::.|:
Db  411 HIVQILVIAGFVSCTMGPLFLALKKGLLRISAEDEWAGMDLTRHGGFAYVYHDDDEHDK 470

Qy  482 ---GMQLR----RVGPNASS 494
            ||:||    ||.| :
Db  471 SVGGFMLRSAQTRVEPAAAA 490
```

RESULT 5

AAO26534
ID   AAO26534 standard; protein; 494 AA.
XX
AC   AAO26534;
XX
DT   18-FEB-2003 (first entry)
XX
DE   Wheat high affinity ammonium transporter protein.
XX
KW   Herbicide; ammonium transporter protein; herbicide; transgenic plant;
KW   wheat.
XX
OS   Triticum aestivum.
XX

```
RESULT 6
AAG40621
ID   AAG40621 standard; protein; 446 AA.
XX
AC   AAG40621;
XX
DT   18-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 50427.
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
XX
CS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-00301439.
XX
PR   25-FEB-1999; 99US-0121825P.   PR   18-JUN-1999; 99US-0139458P.
PR   05-MAR-1999; 99US-0123180P.   PR   18-JUN-1999; 99US-0139459P.
PR   09-MAR-1999; 99US-0123548P.   PR   18-JUN-1999; 99US-0139460P.
PR   23-MAR-1999; 99US-0125788P.   PR   18-JUN-1999; 99US-0139461P.
PR   25-MAR-1999; 99US-0126264P.   PR   18-JUN-1999; 99US-0139462P.
PR   29-MAR-1999; 99US-0126785P.   PR   18-JUN-1999; 99US-0139463P.
PR   01-APR-1999; 99US-0127462P.   PR   18-JUN-1999; 99US-0139750P.
PR   06-APR-1999; 99US-0128234P.   PR   18-JUN-1999; 99US-0139763P.
PR   08-APR-1999; 99US-0128714P.   PR   21-JUN-1999; 99US-0139817P.
PR   16-APR-1999; 99US-0129845P.   PR   22-JUN-1999; 99US-0139899P.
PR   19-APR-1999; 99US-0130077P.   PR   23-JUN-1999; 99US-0140353P.
PR   21-APR-1999; 99US-0130449P.   PR   23-JUN-1999; 99US-0140354P.
PR   23-APR-1999; 99US-0130510P.   PR   24-JUN-1999; 99US-0140695P.
PR   23-APR-1999; 99US-0130891P.   PR   28-JUN-1999; 99US-0140823P.
PR   28-APR-1999; 99US-0131449P.   PR   29-JUN-1999; 99US-0140991P.
PR   30-APR-1999; 99US-0132048P.   PR   30-JUN-1999; 99US-0141287P.
PR   30-APR-1999; 99US-0132407P.   PR   01-JUL-1999; 99US-0141842P.
PR   04-MAY-1999; 99US-0132484P.   PR   01-JUL-1999; 99US-0142154P.
PR   05-MAY-1999; 99US-0132485P.   PR   02-JUL-1999; 99US-0142055P.
PR   06-MAY-1999; 99US-0132486P.   PR   06-JUL-1999; 99US-0142390P.
PR   06-MAY-1999; 99US-0132487P.   PR   08-JUL-1999; 99US-0142803P.
PR   07-MAY-1999; 99US-0132863P.   PR   09-JUL-1999; 99US-0142920P.
PR   11-MAY-1999; 99US-0134256P.   PR   09-JUL-1999; 99US-0142977P.
PR   14-MAY-1999; 99US-0134218P.   PR   12-JUL-1999; 99US-0143542P.
PR   14-MAY-1999; 99US-0134219P.   PR   13-JUL-1999; 99US-0143624P.
PR   14-MAY-1999; 99US-0134221P.   PR   14-JUL-1999; 99US-0144005P.
PR   14-MAY-1999; 99US-0134370P.   PR   15-JUL-1999; 99US-0144085P.
PR   18-MAY-1999; 99US-0134768P.   PR   16-JUL-1999; 99US-0144086P.
PR   19-MAY-1999; 99US-0134941P.   PR   19-JUL-1999; 99US-0144325P.
PR   20-MAY-1999; 99US-0135124P.   PR   19-JUL-1999; 99US-0144331P.
PR   21-MAY-1999; 99US-0135353P.   PR   19-JUL-1999; 99US-0144332P.
PR   24-MAY-1999; 99US-0135629P.   PR   19-JUL-1999; 99US-0144333P.
PR   25-MAY-1999; 99US-0136021P.   PR   19-JUL-1999; 99US-0144334P.
PR   27-MAY-1999; 99US-0136392P.   PR   19-JUL-1999; 99US-0144335P.
PR   28-MAY-1999; 99US-0137502P.   PR   20-JUL-1999; 99US-0144352P.
PR   01-JUN-1999; 99US-0137222P.   PR   20-JUL-1999; 99US-0144632P.
PR   03-JUN-1999; 99US-0137528P.   PR   20-JUL-1999; 99US-0144884P.
PR   04-JUN-1999; 99US-0137502P.   PR   21-JUL-1999; 99US-0144814P.
PR   07-JUN-1999; 99US-0137724P.   PR   21-JUL-1999; 99US-0145086P.
PR   08-JUN-1999; 99US-0138094P.   PR   22-JUL-1999; 99US-0145088P.
PR   10-JUN-1999; 99US-0138540P.   PR   22-JUL-1999; 99US-0145085P.
PR   10-JUN-1999; 99US-0138847P.   PR   22-JUL-1999; 99US-0145087P.
PR   14-JUN-1999; 99US-0139119P.   PR   22-JUL-1999; 99US-0145089P.
PR   16-JUN-1999; 99US-0139452P.   PR   22-JUL-1999; 99US-0145192P.
PR   16-JUN-1999; 99US-0139453P.   PR   23-JUL-1999; 99US-0145145P.
PR   17-JUN-1999; 99US-0139492P.   PR   23-JUL-1999; 99US-0145218P.
PR   18-JUN-1999; 99US-0139454P.   PR   23-JUL-1999; 99US-0145224P.
PR   18-JUN-1999; 99US-0139455P.   PR   26-JUL-1999; 99US-0145276P.
PR   18-JUN-1999; 99US-0139456P.   PR   27-JUL-1999; 99US-0145913P.
PR   18-JUN-1999; 99US-0139457P.   PR   27-JUL-1999; 99US-0145518P.
                                   PR   27-JUL-1999; 99US-0145919P.
                                   PR   28-JUL-1999; 99US-0145951P.
                                   PR   02-AUG-1999; 99US-0146386P.
                                   PR   02-AUG-1999; 99US-0146388P.
                                   PR   03-AUG-1999; 99US-0146389P.
                                   PR   04-AUG-1999; 99US-0147038P.
                                   PR   04-AUG-1999; 99US-0147204P.
                                   PR   05-AUG-1999; 99US-0147302P.
                                   PR   05-AUG-1999; 99US-0147192P.
                                   PR   06-AUG-1999; 99US-0147260P.
                                   PR   06-AUG-1999; 99US-0147303P.
                                   PR   09-AUG-1999; 99US-0147416P.
                                   PR   09-AUG-1999; 99US-0147493P.
                                   PR   10-AUG-1999; 99US-0147935P.
                                   PR   11-AUG-1999; 99US-0148171P.
                                   PR   12-AUG-1999; 99US-0148319P.
                                   PR   13-AUG-1999; 99US-0148341P.
                                   PR   13-AUG-1999; 99US-0148565P.
                                   PR   16-AUG-1999; 99US-0148684P.
                                   PR   17-AUG-1999; 99US-0149368P.
                                   PR   18-AUG-1999; 99US-0149175P.
                                   PR   18-AUG-1999; 99US-0149426P.
                                   PR   20-AUG-1999; 99US-0149722P.
```

```
Qy  241  LRGHSASLVVLGTFLLWFGWYGFNPGSFNKILLTYGNSGNYYGQWSAVGRTAVTTTLAGS  300
         ||||||||||||||||||||  ||||||||||||:| : :|| ||:||||||||||||||
Db  181  LRGHSASLVVLGTFLLWFGFYGFNPGSFNKILVTY-ETGTYNGQWSAVGRTAVTTTLAGC  239

Qy  301  TAALTTLFGKRVISGHWNVTDVCNGLLGGFAAITAGCSVVFPWAAIVCGFVASIVLIACN  360
         ||||||||||| :|||||||||||||||||||||| |||||||||| :||||| :||| |
Db  240  TAALTTLFGKRLLSGHWNVTDVCNGLLGGFAAITGGCSVVFPWAAIICGFVAALVLLGCN  299

Qy  361  KLAEKVKFDDPLEAAQLHGGCGTWGVIFTALFAKKEYVKEYVG--LGRAHGLLMGGGGKL  418
         ||||| |:||||| ||||||||: ||:||||||::: :: :|      | ||:|||||||
Db  300  KLAEKLKYDDPLERAQLHGGCGANGLIFTALFAQEKYLNQIYGNKPGRPHGLFMGGGGKL  359

Qy  419  LAAHVIQILVIAGWVSATMGPLFWGLNKLKLLRISSEDELAGMDMTRHGGFAYAY-EDDE  477
         |  :|||::|| |||||||| ||: |||:| |||||||| |||||||||||||:| :|||
Db  360  LGAQLIQIVITGWVSATMGTLFFILKKMKLLRISSEDEMAGMDMTRHGGFAYMYFDDDE   419

Qy  478  THKHGMQLRRVGP-----NASSTPT  497
         :||  :|||||      | :: |||
Db  420  SHK-AIQLRRVEPRSPSPSGANTTPT  444
```

RESULT 7
AAG36234
ID   AAG36234 standard; protein; 455 AA.
XX
AC   AAG36234;
XX
DT   18-OCT-2000 (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 44376.
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-00301439.
XX

```
PR  20-AUG-1999;   99US-0149723P.        PR  25-FEB-1999;   99US-0121825P.
PR  20-AUG-1999;   99US-0149929P.        PR  05-MAR-1999;   99US-0123180P.
PR  23-AUG-1999;   99US-0149902P.        PR  09-MAR-1999;   99US-0123548P.
PR  23-AUG-1999;   99US-0149930P.        PR  23-MAR-1999;   99US-0125788P.
PR  25-AUG-1999;   99US-0150566P.        PR  25-MAR-1999;   99US-0126264P.
PR  26-AUG-1999;   99US-0150884P.        PR  29-MAR-1999;   99US-0126785P.
PR  26-AUG-1999;   99US-0151065P.        PR  01-APR-1999;   99US-0127462P.
PR  27-AUG-1999;   99US-0151066P.        PR  06-APR-1999;   99US-0128234P.
PR  27-AUG-1999;   99US-0151080P.        PR  08-APR-1999;   99US-0128714P.
PR  30-AUG-1999;   99US-0151303P.        PR  16-APR-1999;   99US-0129845P.
PR  31-AUG-1999;   99US-0151438P.        PR  19-APR-1999;   99US-0130077P.
PR  01-SEP-1999;   99US-0151930P.        PR  21-APR-1999;   99US-0130449P.
PR  07-SEP-1999;   99US-0152363P.        PR  23-APR-1999;   99US-0130510P.
PR  10-SEP-1999;   99US-0153070P.        PR  23-APR-1999;   99US-0130891P.
PR  13-SEP-1999;   99US-0153758P.        PR  28-APR-1999;   99US-0131449P.
PR  15-SEP-1999;   99US-0154018P.        PR  30-APR-1999;   99US-0132048P.
PR  16-SEP-1999;   99US-0154039P.        PR  30-APR-1999;   99US-0132407P.
PR  20-SEP-1999;   99US-0154779P.        PR  04-MAY-1999;   99US-0132484P.
PR  22-SEP-1999;   99US-0155139P.        PR  05-MAY-1999;   99US-0132485P.
PR  23-SEP-1999;   99US-0155486P.        PR  06-MAY-1999;   99US-0132486P.
PR  24-SEP-1999;   99US-0155659P.        PR  06-MAY-1999;   99US-0132487P.
PR  28-SEP-1999;   99US-0156458P.        PR  07-MAY-1999;   99US-0132863P.
PR  29-SEP-1999;   99US-0156596P.        PR  11-MAY-1999;   99US-0134256P.
PR  04-OCT-1999;   99US-0157117P.        PR  14-MAY-1999;   99US-0134218P.
PR  05-OCT-1999;   99US-0157753P.        PR  14-MAY-1999;   99US-0134219P.
PR  06-OCT-1999;   99US-0157865P.        PR  14-MAY-1999;   99US-0134221P.
PR  07-OCT-1999;   99US-0158029P.        PR  14-MAY-1999;   99US-0134370P.
PR  08-OCT-1999;   99US-0158322P.        PR  18-MAY-1999;   99US-0134768P.
PR  12-OCT-1999;   99US-0158369P.        PR  19-MAY-1999;   99US-0134941P.
PR  13-OCT-1999;   99US-0159293P.
PR  13-OCT-1999;   99US-0159294P.
PR  13-OCT-1999;   99US-0159295P.
PR  14-OCT-1999;   99US-0159329P.
PR  14-OCT-1999;   99US-0159330P.
PR  14-OCT-1999;   99US-0159331P.
PR  14-OCT-1999;   99US-0159637P.
PR  14-OCT-1999;   99US-0159638P.
PR  18-OCT-1999;   99US-0159584P.
PR  21-OCT-1999;   99US-0160741P.
PR  21-OCT-1999;   99US-0160767P.
PR  21-OCT-1999;   99US-0160768P.
PR  21-OCT-1999;   99US-0160770P.
PR  21-OCT-1999;   99US-0160814P.
PR  21-OCT-1999;   99US-0160815P.
PR  22-OCT-1999;   99US-0160980P.
PR  22-OCT-1999;   99US-0160981P.
PR  22-OCT-1999;   99US-0160989P.
PR  25-OCT-1999;   99US-0161404P.
PR  25-OCT-1999;   99US-0161405P.
PR  25-OCT-1999;   99US-0161406P.
PR  26-OCT-1999;   99US-0161359P.
PR  26-OCT-1999;   99US-0161360P.
PR  26-OCT-1999;   99US-0161361P.
PR  28-OCT-1999;   99US-0161920P.
PR  28-OCT-1999;   99US-0161992P.
PR  28-OCT-1999;   99US-0161993P.
PR  29-OCT-1999;   99US-0162142P.
```

```
Query Match         73.6%;  Score 1943;  DB 3;  Length 446;
Best Local Similarity   82.7%;
Matches  369;  Conservative  35;  Mismatches  30;  Indels  12;  Gaps  6;   Pred. No. 2.2e-183;

Qy   62  MQLGFAMLCAGSVRAKNTMNIMLTNVLDAAAAGGLFYYLFGPAPAFGSPSNGFIGKHFFGL  121
         |||||||||||||||||||||||||||||||||||||||||| |||||||||||||||:||
Db    1  MQLGFAMLCAGSVRAKNTMNIMLTNVLDAAAAGGLFYYLFGYAFAFGSPSNGFIGKHYFGL   60

Qy  122  KDIPSSSYDYSYFLYQWAFAIAAAGITSGSIAERTQFVAYLIYSSFLTGFVYPVVSHWFW  181
         ||||: : :||:|||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  KDIPTASADYSNFLYQWAFAIAAAGITSGSIAERTQFVAYLIYSSFLTGFVYPVVSHWFW  120

Qy  182  SPDGWASAFKIT-DRLFSTGVIDFAGSGVVHMVGGIAGLWGALIEGPRMGRFDHAGRAVA  240
         | |||||: | :|||||||:|||||||||||||||||||||||||||:||||    ||:|
Db  121  SVDGWASPFRTDGDLLFSTGAIDFAGSGVVHMVGGIAGLWGALIEGPRLGRFDNGGRAIA  180
```

```
PR  20-MAY-1999;  99US-0135124P.
PR  21-MAY-1999;  99US-0135353P.
PR  24-MAY-1999;  99US-0135629P.
PR  25-MAY-1999;  99US-0136021P.
PR  27-MAY-1999;  99US-0136392P.
PR  28-MAY-1999;  99US-0136782P.
PR  01-JUN-1999;  99US-0137222P.
PR  03-JUN-1999;  99US-0137528P.
PR  04-JUN-1999;  99US-0137502P.
PR  07-JUN-1999;  99US-0137724P.
PR  08-JUN-1999;  99US-0138094P.
PR  10-JUN-1999;  99US-0138540P.
PR  10-JUN-1999;  99US-0138847P.
PR  14-JUN-1999;  99US-0139119P.
PR  16-JUN-1999;  99US-0139452P.
PR  16-JUN-1999;  99US-0139453P.
PR  17-JUN-1999;  99US-0139492P.
PR  18-JUN-1999;  99US-0139454P.
PR  18-JUN-1999;  99US-0139455P.
PR  18-JUN-1999;  99US-0139456P.
PR  18-JUN-1999;  99US-0139457P.
PR  18-JUN-1999;  99US-0139458P.
PR  18-JUN-1999;  99US-0139459P.
PR  18-JUN-1999;  99US-0139460P.
PR  18-JUN-1999;  99US-0139461P.
PR  18-JUN-1999;  99US-0139462P.
PR  18-JUN-1999;  99US-0139463P.
PR  18-JUN-1999;  99US-0139750P.
PR  18-JUN-1999;  99US-0139763P.
PR  21-JUN-1999;  99US-0139817P.
PR  22-JUN-1999;  99US-0139899P.
PR  23-JUN-1999;  99US-0140353P.
PR  23-JUN-1999;  99US-0140354P.
PR  24-JUN-1999;  99US-0140695P.
PR  28-JUN-1999;  99US-0140823P.
PR  29-JUN-1999;  99US-0140991P.
PR  30-JUN-1999;  99US-0141287P.
PR  01-JUL-1999;  99US-0141842P.
PR  01-JUL-1999;  99US-0142154P.
PR  02-JUL-1999;  99US-0142055P.
PR  06-JUL-1999;  99US-0142390P.
PR  08-JUL-1999;  99US-0142803P.
PR  09-JUL-1999;  99US-0142920P.
PR  12-JUL-1999;  99US-0142977P.
PR  13-JUL-1999;  99US-0143542P.
PR  14-JUL-1999;  99US-0143624P.
PR  15-JUL-1999;  99US-0144005P.
PR  16-JUL-1999;  99US-0144085P.
PR  16-JUL-1999;  99US-0144086P.
PR  19-JUL-1999;  99US-0144325P.
PR  19-JUL-1999;  99US-0144331P.
PR  19-JUL-1999;  99US-0144332P.
PR  19-JUL-1999;  99US-0144333P.
PR  19-JUL-1999;  99US-0144334P.
PR  19-JUL-1999;  99US-0144335P.
PR  20-JUL-1999;  99US-0144352P.
PR  20-JUL-1999;  99US-0144632P.
PR  21-JUL-1999;  99US-0144884P.
PR  21-JUL-1999;  99US-0144814P.
PR  21-JUL-1999;  99US-0145086P.
PR  21-JUL-1999;  99US-0145088P.
PR  22-JUL-1999;  99US-0145085P.
PR  22-JUL-1999;  99US-0145087P.
PR  22-JUL-1999;  99US-0145089P.
PR  22-JUL-1999;  99US-0145192P.
PR  23-JUL-1999;  99US-0145145P.
PR  23-JUL-1999;  99US-0145218P.
PR  23-JUL-1999;  99US-0145224P.
PR  26-JUL-1999;  99US-0145276P.
PR  27-JUL-1999;  99US-0145913P.
PR  27-JUL-1999;  99US-0145918P.
PR  27-JUL-1999;  99US-0145919P.
PR  28-JUL-1999;  99US-0145951P.

PR  02-AUG-1999;  99US-0146386P.
PR  02-AUG-1999;  99US-0146388P.
PR  02-AUG-1999;  99US-0146389P.
PR  03-AUG-1999;  99US-0147038P.
PR  04-AUG-1999;  99US-0147204P.
PR  04-AUG-1999;  99US-0147302P.
PR  05-AUG-1999;  99US-0147192P.
PR  05-AUG-1999;  99US-0147260P.
PR  06-AUG-1999;  99US-0147303P.
PR  06-AUG-1999;  99US-0147416P.
PR  09-AUG-1999;  99US-0147493P.
PR  09-AUG-1999;  99US-0147935P.
PR  10-AUG-1999;  99US-0148171P.
PR  11-AUG-1999;  99US-0148319P.
PR  12-AUG-1999;  99US-0148341P.
PR  13-AUG-1999;  99US-0148565P.
PR  13-AUG-1999;  99US-0148684P.
PR  16-AUG-1999;  99US-0149368P.
PR  17-AUG-1999;  99US-0149175P.
PR  18-AUG-1999;  99US-0149426P.
PR  20-AUG-1999;  99US-0149722P.
PR  20-AUG-1999;  99US-0149723P.
PR  20-AUG-1999;  99US-0149929P.
PR  23-AUG-1999;  99US-0149902P.
PR  23-AUG-1999;  99US-0149930P.
PR  25-AUG-1999;  99US-0150566P.
PR  26-AUG-1999;  99US-0150584P.
PR  27-AUG-1999;  99US-0151065P.
PR  27-AUG-1999;  99US-0151066P.
PR  30-AUG-1999;  99US-0151080P.
PR  31-AUG-1999;  99US-0151303P.
PR  31-AUG-1999;  99US-0151438P.
PR  01-SEP-1999;  99US-0151930P.
PR  07-SEP-1999;  99US-0152363P.
PR  10-SEP-1999;  99US-0153070P.
PR  13-SEP-1999;  99US-0153758P.
PR  15-SEP-1999;  99US-0154018P.
PR  16-SEP-1999;  99US-0154039P.
PR  20-SEP-1999;  99US-0154779P.
PR  22-SEP-1999;  99US-0155139P.
PR  23-SEP-1999;  99US-0155486P.
PR  24-SEP-1999;  99US-0155659P.
PR  28-SEP-1999;  99US-0156458P.
PR  29-SEP-1999;  99US-0155596P.
PR  04-OCT-1999;  99US-0157117P.
PR  06-OCT-1999;  99US-0157753P.
PR  06-OCT-1999;  99US-0157865P.
PR  07-OCT-1999;  99US-0158029P.
PR  08-OCT-1999;  99US-0158232P.
PR  12-OCT-1999;  99US-0158369P.
PR  13-OCT-1999;  99US-0159293P.
PR  13-OCT-1999;  99US-0159294P.
PR  13-OCT-1999;  99US-0159295P.
PR  14-OCT-1999;  99US-0159329P.
PR  14-OCT-1999;  99US-0159330P.
PR  14-OCT-1999;  99US-0159331P.
PR  18-OCT-1999;  99US-0159584P.
PR  21-OCT-1999;  99US-0160741P.
PR  21-OCT-1999;  99US-0160767P.
PR  21-OCT-1999;  99US-0160768P.
PR  21-OCT-1999;  99US-0160770P.
PR  21-OCT-1999;  99US-0160814P.
PR  22-OCT-1999;  99US-0160815P.
PR  22-OCT-1999;  99US-0160980P.
PR  22-OCT-1999;  99US-0160981P.
PR  25-OCT-1999;  99US-0160989P.
PR  25-OCT-1999;  99US-0161404P.
PR  25-OCT-1999;  99US-0161405P.
PR  25-OCT-1999;  99US-0161406P.
PR  26-OCT-1999;  99US-0161359P.
PR  26-OCT-1999;  99US-0161360P.
```

```
PR   26-OCT-1999;  99US-0161361P.
PR   28-OCT-1999;  99US-0161920P.
PR   28-OCT-1999;  99US-0161992P.
PR   28-OCT-1999;  99US-0161993P.
PR   29-OCT-1999;  99US-0162142P.

Query Match     73.6%;  Score 1943;  DB 3;  Length 455;
Best Local Similarity  82.7%;  Pred. No. 2.2e-183;
Matches 369;  Conservative 35;  Mismatches 30;  Indels 12;  Gaps 6;

Qy   62  MQLGFAMLCAGSVRAKNTMNIMLTNVLDAAAGGLFYYLFGFAFAFGSPSNGFIGKHFFGL 121
         |||||||||||||||||||||||||||||||||||||||::||||||||||||||||:||
Db    1  MQLGFAMLCAGSVRAKNTMNIMLTNVLDAAAGGLFYYLFGYAFAFGSPSNGFIGKHYFGL 60

Qy  122  KDIPSSSYDYSYFLYQWAFAIAAAGITSGSIAERTQFVAYLIYSSFLTGFVYPVVSHWFW 181
         :|||:::    ||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  KDIPTASADYISNFLIYQWAFAIAAAGITSGSIAERTQFVAYLIYSSFLTGFVYPVVSHWFW 120

Qy  182  SPDGWASAFKIT-DRLFSTGVIDFAGSGVVHMVGGIAGLWGALIEGPRMGRFDHAGRAVA 240
         | ||||| :    |||||| :|||||||||||||||||||||||||||:||||| ||:|
Db  121  SVDGWASPFRTDGDLLFSTGAIDFAGSGVVHMVGGIAGLWGALIEGPRLGRFDNGGRAIA 180

Qy  241  LRGHSASLVVLGTFLLWFGWYGFNPGSFNKILLTYGNSGNYYGQWSAVGRTAVTTLLAGS 300
         ||||||||||||||||||||||||||||||||:||  :|  |:||||||||||||||||
Db  181  LRGHSASLVVLGTFLLWFGWYGFNPGSFNKILVTY-ETGTYNGQWSAVGRTAVTTLLAGC 239

Qy  301  TAALTTLFGKRVISGHWNVTDVCNGLLGGFAAITAGCSVVEPWAAIVCGFVASIVLIACN 360
         |||||||||||:  ||||||||||||||||||:|:||||||::|||:|||||||:||
Db  240  TAALTTLFGKRLLSGHWNVTDVCNGLLGGFAAITGGCSVVEPWAAIICGFVAALVLLGCN 299

Qy  361  KLABKVKFDDPLEAAQLHGGCGTWGVIFTALFAKKEYVKEVVG--LGRAHGLLMGGGGKL 418
         ||:|||||||||||||||||||:||||||||:: :  :|||    |||:|||:|||||||
Db  300  KLABKLKYDDPLEAAQLHGGCGAWGLIFTALFAQEKYLNQIYGNKPGRPHGLFMGGGGKL 359

Qy  419  LAAHVIQILVIAGWVSATMGPLFWGLNKLKLLRISSEDELAGMDMTRHGGFAYAY-EDDE 477
         | |::|||| ||:|||||||:||  |:|:|||||||||||:|||||||||||||  :||
Db  360  LGAQLIQIIVITGWVSATMGTLFFILKKMKLRISSEDEMAGMDMTRHGGFAYMYFDDDE 419

Qy  478  THKHGMQLRRVGP------NASSTPT 497
         :||:|:|||||       |::|||
Db  420  SHK-AIQLRRVEPRSPSPSGANTTPT 444
```

RESULT 8
AAG40622
ID   AAG40622 standard; protein; 440 AA.
XX
AC   AAG40622;
XX
DT   18-OCT-2000 (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 50428.
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-00301439.
XX

```
PR   25-FEB-1999;  99US-0121825P.
PR   05-MAR-1999;  99US-0123180P.
PR   09-MAR-1999;  99US-0123548P.
PR   23-MAR-1999;  99US-0125788P.
PR   25-MAR-1999;  99US-0126264P.
PR   29-MAR-1999;  99US-0126785P.
PR   01-APR-1999;  99US-0127462P.
PR   06-APR-1999;  99US-0128234P.

PR   08-APR-1999;  99US-0128714P.
PR   16-APR-1999;  99US-0129845P.
PR   19-APR-1999;  99US-0130077P.
PR   21-APR-1999;  99US-0130449P.
PR   23-APR-1999;  99US-0130510P.
PR   28-APR-1999;  99US-0130881P.
PR   28-APR-1999;  99US-0131449P.
PR   30-APR-1999;  99US-0132048P.
PR   30-APR-1999;  99US-0132407P.
PR   04-MAY-1999;  99US-0132484P.
PR   05-MAY-1999;  99US-0132485P.
PR   06-MAY-1999;  99US-0132486P.
PR   06-MAY-1999;  99US-0132487P.
PR   07-MAY-1999;  99US-0132863P.
PR   11-MAY-1999;  99US-0134256P.
PR   14-MAY-1999;  99US-0134218P.
PR   14-MAY-1999;  99US-0134219P.
PR   14-MAY-1999;  99US-0134221P.
PR   14-MAY-1999;  99US-0134370P.
PR   18-MAY-1999;  99US-0134768P.
PR   19-MAY-1999;  99US-0134941P.
PR   20-MAY-1999;  99US-0135124P.
PR   21-MAY-1999;  99US-0135353P.
PR   24-MAY-1999;  99US-0135629P.
PR   25-MAY-1999;  99US-0136021P.
PR   27-MAY-1999;  99US-0136392P.
PR   28-MAY-1999;  99US-0136782P.
PR   01-JUN-1999;  99US-0137222P.
PR   03-JUN-1999;  99US-0137528P.
PR   04-JUN-1999;  99US-0137502P.
PR   07-JUN-1999;  99US-0137724P.
PR   08-JUN-1999;  99US-0138094P.
PR   10-JUN-1999;  99US-0138540P.
PR   10-JUN-1999;  99US-0138847P.
PR   14-JUN-1999;  99US-0139119P.
PR   16-JUN-1999;  99US-0139452P.
PR   16-JUN-1999;  99US-0139453P.
PR   17-JUN-1999;  99US-0139492P.
PR   18-JUN-1999;  99US-0139454P.
PR   18-JUN-1999;  99US-0139455P.
PR   18-JUN-1999;  99US-0139456P.
PR   18-JUN-1999;  99US-0139457P.
PR   18-JUN-1999;  99US-0139458P.
PR   18-JUN-1999;  99US-0139459P.
PR   18-JUN-1999;  99US-0139460P.
PR   18-JUN-1999;  99US-0139461P.
PR   18-JUN-1999;  99US-0139462P.
PR   18-JUN-1999;  99US-0139463P.
PR   18-JUN-1999;  99US-0139750P.
PR   18-JUN-1999;  99US-0139763P.
PR   21-JUN-1999;  99US-0139817P.
PR   22-JUN-1999;  99US-0139899P.
PR   23-JUN-1999;  99US-0140353P.
PR   23-JUN-1999;  99US-0140354P.
PR   24-JUN-1999;  99US-0140695P.
PR   28-JUN-1999;  99US-0140823P.
PR   29-JUN-1999;  99US-0140991P.
PR   30-JUN-1999;  99US-0141287P.
PR   01-JUL-1999;  99US-0141842P.
PR   01-JUL-1999;  99US-0142154P.
PR   02-JUL-1999;  99US-0142055P.
PR   06-JUL-1999;  99US-0142390P.
PR   08-JUL-1999;  99US-0142803P.
PR   09-JUL-1999;  99US-0142920P.
PR   12-JUL-1999;  99US-0142977P.
PR   13-JUL-1999;  99US-0143542P.
PR   14-JUL-1999;  99US-0143624P.
PR   15-JUL-1999;  99US-0144005P.
PR   16-JUL-1999;  99US-0144085P.
PR   16-JUL-1999;  99US-0144086P.
PR   19-JUL-1999;  99US-0144325P.
PR   19-JUL-1999;  99US-0144331P.
PR   19-JUL-1999;  99US-0144332P.
```

```
PR   99US-0144333P.   19-JUL-1999;
PR   99US-0144334P.   19-JUL-1999;
PR   99US-0144335P.   19-JUL-1999;
PR   99US-0144352P.   20-JUL-1999;
PR   99US-0144632P.   20-JUL-1999;
PR   99US-0144884P.   21-JUL-1999;
PR   99US-0144814P.   21-JUL-1999;
PR   99US-0145086P.   21-JUL-1999;
PR   99US-0145088P.   21-JUL-1999;
PR   99US-0145085P.   22-JUL-1999;
PR   99US-0145087P.   22-JUL-1999;
PR   99US-0145089P.   22-JUL-1999;
PR   99US-0145192P.   22-JUL-1999;
PR   99US-0145145P.   23-JUL-1999;
PR   99US-0145218P.   23-JUL-1999;
PR   99US-0145224P.   23-JUL-1999;
PR   99US-0145276P.   26-JUL-1999;
PR   99US-0145913P.   27-JUL-1999;
PR   99US-0145918P.   27-JUL-1999;
PR   99US-0145919P.   27-JUL-1999;
PR   99US-0145951P.   28-JUL-1999;
PR   99US-0146386P.   02-AUG-1999;
PR   99US-0146388P.   02-AUG-1999;
PR   99US-0146389P.   02-AUG-1999;
PR   99US-0147038P.   03-AUG-1999;
PR   99US-0147204P.   04-AUG-1999;
PR   99US-0147302P.   04-AUG-1999;
PR   99US-0147192P.   05-AUG-1999;
PR   99US-0147260P.   05-AUG-1999;
PR   99US-0147416P.   06-AUG-1999;
PR   99US-0147493P.   09-AUG-1999;
PR   99US-0147935P.   09-AUG-1999;
PR   99US-0148171P.   10-AUG-1999;
PR   99US-0148319P.   11-AUG-1999;
PR   99US-0148341P.   12-AUG-1999;
PR   99US-0148565P.   13-AUG-1999;
PR   99US-0148684P.   13-AUG-1999;
PR   99US-0149368P.   16-AUG-1999;
PR   99US-0149175P.   17-AUG-1999;
PR   99US-0149426P.   18-AUG-1999;
PR   99US-0149722P.   20-AUG-1999;
PR   99US-0149723P.   20-AUG-1999;
PR   99US-0149929P.   20-AUG-1999;
PR   99US-0149902P.   23-AUG-1999;
PR   99US-0149930P.   25-AUG-1999;
PR   99US-0150566P.   25-AUG-1999;
PR   99US-0150884P.   26-AUG-1999;
PR   99US-0151065P.   27-AUG-1999;
PR   99US-0151066P.   27-AUG-1999;
PR   99US-0151080P.   30-AUG-1999;
PR   99US-0151303P.   31-AUG-1999;
PR   99US-0151438P.   01-SEP-1999;
PR   99US-0151930P.   07-SEP-1999;
PR   99US-0152363P.   10-SEP-1999;
PR   99US-0153070P.   11-SEP-1999;
PR   99US-0153758P.   15-SEP-1999;
PR   99US-0154018P.   16-SEP-1999;
PR   99US-0154039P.   20-SEP-1999;
PR   99US-0154779P.   22-SEP-1999;
PR   99US-0155139P.   23-SEP-1999;
PR   99US-0155486P.   24-SEP-1999;
PR   99US-0155659P.   28-SEP-1999;
PR   99US-0156458P.   28-SEP-1999;
PR   99US-0156596P.   29-SEP-1999;
PR   99US-0157117P.   04-OCT-1999;
PR   99US-0157753P.   05-OCT-1999;
PR   99US-0157865P.   06-OCT-1999;
PR   99US-0158029P.   07-OCT-1999;
PR   99US-0158232P.   08-OCT-1999;
PR   99US-0158369P.   12-OCT-1999;
PR   99US-0159293P.   13-OCT-1999;
PR   99US-0159294P.   13-OCT-1999;

PR   99US-0159295P.   13-OCT-1999;
PR   99US-0159329P.   14-OCT-1999;
PR   99US-0159330P.   14-OCT-1999;
PR   99US-0159331P.   14-OCT-1999;
PR   99US-0159637P.   14-OCT-1999;
PR   99US-0159638P.   14-OCT-1999;
PR   99US-0159584P.   18-OCT-1999;
PR   99US-0160741P.   21-OCT-1999;
PR   99US-0160767P.   21-OCT-1999;
PR   99US-0160768P.   21-OCT-1999;
PR   99US-0160770P.   21-OCT-1999;
PR   99US-0160814P.   21-OCT-1999;
PR   99US-0160815P.   21-OCT-1999;
PR   99US-0160980P.   22-OCT-1999;
PR   99US-0160981P.   22-OCT-1999;
PR   99US-0160989P.   22-OCT-1999;
PR   99US-0161404P.   25-OCT-1999;
PR   99US-0161405P.   25-OCT-1999;
PR   99US-0161406P.   25-OCT-1999;
PR   99US-0161359P.   26-OCT-1999;
PR   99US-0161360P.   26-OCT-1999;
PR   99US-0161361P.   26-OCT-1999;
PR   99US-0161920P.   28-OCT-1999;
PR   99US-0161992P.   28-OCT-1999;
PR   99US-0161993P.   28-OCT-1999;
PR   99US-0162142P.   29-OCT-1999;
```

```
Query Match            72.5%;      Score 1913;  DB 3;  Length 440;
Best Local Similarity  82.5%;      Pred. No. 2e-180;
Matches  363;  Conservative  35;  Mismatches  30;  Indels  12;  Gaps  6;

Qy    68 MLCAGSVRAKNTMNIMLTNVLDAAAGGLFYYLFGFAFAFGSPSNGFIGKHFFGLKDIPSS 127
         |||||||||||||||||||||||||||||||||||||||||||||||||::||||||
Db     1 MLCAGSVRAKNTMNIMLTNVLDAAAGGLFYYLFGFAFAFGSPSNGFIGKHYFGLKDIPTA 60

Qy   128 SYDYSYFLYQWAFAIAAAGITSGSIAERTQFVAYLIYSSFLTGFVYPVVSHWFWSPDGWA 187
         | ||| |||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db    61 SADYSNFLYQWAFAIAAAGITSGSIAERTQFVAYLIYSSFLTGFVYPVVSHWFWSVDGWA 120

Qy   188 SAFKIT-DRLFSTGVIDFAGSGVVHMVGGIAGLWGALIEGPRMGRFDHAGRAVALRGHSA 246
         |      |  |||||| ||||||||||||||||||||||||: ||||::| |:|||||||
Db   121 SPFRTDGDLLFSTGAIDFAGSGVVHMVGGIAGLWGALIEGPRILGRFDNGGRAIALRGHSA 180

Qy   247 SLVVLGTFLLWFGWYGFNPGSFNKILLTYGNSGNYYGQWSAVGRTAVTTTLAGSTAALTT 306
         |||||||||||||||||||||||||||:||  ||  ||||||||||||||||||:||||| 
Db   181 SLVVLGTFLLWFGWYGFNPGSFNKILVTY-ETGTYNGQWSAVGRTAVTTTLAGCTAALTT 239

Qy   307 LFGKRVISGHWNVTDVCNGLLGGFAAITAGCSVVEPWAAIVCGFVASIVLIACNKLAEKV 366
         |||||::||||||||||||||||||||  ||::||  |||:| |||::||::||||||| 
Db   240 LFGKRLLSGHWNVTDVCNGLLGGFAAIICGFVAALVLLGCNKLAEKL 299

Qy   367 KFDDPLEAAQLHGGCGTWGVIFTALFAKKEYVKEVYG--LGRAHGLLMGGGGKLLAAHVI 424
         |:||||||||||||||:||:|||||||:::: :|:|| :  | | | :|||||||||| :
Db   300 KYDDPLEAAQLHGGCGAWGLIFTALFAQEKYLNQIYGNKPGRPHGLFMGGGGKLLGAQLI 359

Qy   425 QILVIAGWVSATMGPLFWGLNKLKLRISSEDELAGMDMTRHGGFAYAY-EDDETHKHGM 483
         ||:|| ||||||||:|| :|::|||:|||||||:||||||||||:||| :||||||:||
Db   360 QIIVITGWVSATMGTLFFILKMKLLRISSEDEMAGMDMTRHGGFAYMYFDDDESHK-AI 418

Qy   484 QLRRVGP------NASSTPT 497
         |||||:|       :: || 
Db   419 QLRRVEPRSPSPSGANTTPT 438
```

RESULT 9
AAG36235
ID   AAG36235 standard; protein; 449 AA.
XX
AC   AAG36235;
XX
DT   18-OCT-2000 (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 44377.

```
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-00301439.
XX
PR   25-FEB-1999;  99US-0121825P.
PR   05-MAR-1999;  99US-0123180P.
PR   09-MAR-1999;  99US-0123548P.
PR   23-MAR-1999;  99US-0125788P.
PR   25-MAR-1999;  99US-0126264P.
PR   29-MAR-1999;  99US-0126785P.
PR   01-APR-1999;  99US-0127462P.
PR   06-APR-1999;  99US-0128234P.
PR   08-APR-1999;  99US-0128714P.
PR   16-APR-1999;  99US-0129845P.
PR   19-APR-1999;  99US-0130077P.
PR   21-APR-1999;  99US-0130449P.
PR   23-APR-1999;  99US-0130510P.
PR   23-APR-1999;  99US-0130891P.
PR   28-APR-1999;  99US-0131449P.
PR   30-APR-1999;  99US-0132048P.
PR   04-MAY-1999;  99US-0132407P.
PR   05-MAY-1999;  99US-0132484P.
PR   05-MAY-1999;  99US-0132485P.
PR   06-MAY-1999;  99US-0132486P.
PR   06-MAY-1999;  99US-0132487P.
PR   07-MAY-1999;  99US-0132863P.
PR   11-MAY-1999;  99US-0134256P.
PR   14-MAY-1999;  99US-0134218P.
PR   14-MAY-1999;  99US-0134219P.
PR   14-MAY-1999;  99US-0134221P.
PR   18-MAY-1999;  99US-0134768P.
PR   19-MAY-1999;  99US-0134941P.
PR   20-MAY-1999;  99US-0135124P.
PR   21-MAY-1999;  99US-0135353P.
PR   24-MAY-1999;  99US-0135629P.
PR   25-MAY-1999;  99US-0136021P.
PR   27-MAY-1999;  99US-0136392P.
PR   28-MAY-1999;  99US-0136782P.
PR   01-JUN-1999;  99US-0137222P.
PR   03-JUN-1999;  99US-0137528P.
PR   04-JUN-1999;  99US-0137502P.
PR   07-JUN-1999;  99US-0137724P.
PR   08-JUN-1999;  99US-0138094P.
PR   10-JUN-1999;  99US-0138540P.
PR   10-JUN-1999;  99US-0138847P.
PR   14-JUN-1999;  99US-0139119P.
PR   16-JUN-1999;  99US-0139452P.
PR   16-JUN-1999;  99US-0139453P.
PR   17-JUN-1999;  99US-0139492P.
PR   18-JUN-1999;  99US-0139454P.
PR   18-JUN-1999;  99US-0139455P.
PR   18-JUN-1999;  99US-0139456P.
PR   18-JUN-1999;  99US-0139457P.
PR   18-JUN-1999;  99US-0139458P.
PR   18-JUN-1999;  99US-0139459P.
PR   18-JUN-1999;  99US-0139460P.
PR   18-JUN-1999;  99US-0139461P.
PR   18-JUN-1999;  99US-0139462P.
PR   18-JUN-1999;  99US-0139463P.
PR   18-JUN-1999;  99US-0139750P.
PR   18-JUN-1999;  99US-0139763P.
PR   21-JUN-1999;  99US-0139817P.
PR   22-JUN-1999;  99US-0139899P.
PR   23-JUN-1999;  99US-0140353P.
PR   23-JUN-1999;  99US-0140354P.
PR   24-JUN-1999;  99US-0140695P.
PR   28-JUN-1999;  99US-0140823P.
PR   29-JUN-1999;  99US-0140991P.
PR   30-JUN-1999;  99US-0141287P.
PR   01-JUL-1999;  99US-0141842P.
PR   01-JUL-1999;  99US-0142154P.
PR   02-JUL-1999;  99US-0142055P.
PR   06-JUL-1999;  99US-0142390P.
PR   08-JUL-1999;  99US-0142803P.
PR   09-JUL-1999;  99US-0142920P.
PR   12-JUL-1999;  99US-0142977P.
PR   13-JUL-1999;  99US-0143542P.
PR   14-JUL-1999;  99US-0143624P.
PR   15-JUL-1999;  99US-0144005P.
PR   16-JUL-1999;  99US-0144085P.
PR   16-JUL-1999;  99US-0144086P.
PR   19-JUL-1999;  99US-0144325P.
PR   19-JUL-1999;  99US-0144331P.
PR   19-JUL-1999;  99US-0144332P.
PR   19-JUL-1999;  99US-0144333P.
PR   19-JUL-1999;  99US-0144334P.
PR   19-JUL-1999;  99US-0144335P.
PR   20-JUL-1999;  99US-0144352P.
PR   20-JUL-1999;  99US-0144632P.
PR   21-JUL-1999;  99US-0144884P.
PR   21-JUL-1999;  99US-0144814P.
PR   21-JUL-1999;  99US-0145086P.
PR   22-JUL-1999;  99US-0145088P.
PR   22-JUL-1999;  99US-0145085P.
PR   22-JUL-1999;  99US-0145087P.
PR   22-JUL-1999;  99US-0145089P.
PR   22-JUL-1999;  99US-0145192P.
PR   23-JUL-1999;  99US-0145145P.
PR   23-JUL-1999;  99US-0145218P.
PR   23-JUL-1999;  99US-0145224P.
PR   26-JUL-1999;  99US-0145276P.
PR   27-JUL-1999;  99US-0145913P.
PR   27-JUL-1999;  99US-0145918P.
PR   27-JUL-1999;  99US-0145919P.
PR   28-JUL-1999;  99US-0145951P.
PR   02-AUG-1999;  99US-0146386P.
PR   02-AUG-1999;  99US-0146388P.
PR   03-AUG-1999;  99US-0146389P.
PR   03-AUG-1999;  99US-0147038P.
PR   04-AUG-1999;  99US-0147204P.
PR   04-AUG-1999;  99US-0147302P.
PR   05-AUG-1999;  99US-0147192P.
PR   05-AUG-1999;  99US-0147260P.
PR   06-AUG-1999;  99US-0147303P.
PR   06-AUG-1999;  99US-0147416P.
PR   09-AUG-1999;  99US-0147493P.
PR   09-AUG-1999;  99US-0147935P.
PR   10-AUG-1999;  99US-0148171P.
PR   11-AUG-1999;  99US-0148319P.
PR   12-AUG-1999;  99US-0148341P.
PR   13-AUG-1999;  99US-0148565P.
PR   13-AUG-1999;  99US-0148684P.
PR   17-AUG-1999;  99US-0149368P.
PR   17-AUG-1999;  99US-0149175P.
PR   18-AUG-1999;  99US-0149426P.
PR   20-AUG-1999;  99US-0149722P.
PR   20-AUG-1999;  99US-0149723P.
PR   20-AUG-1999;  99US-0149929P.
PR   23-AUG-1999;  99US-0149902P.
PR   23-AUG-1999;  99US-0149930P.
PR   25-AUG-1999;  99US-0150566P.
PR   26-AUG-1999;  99US-0150884P.
PR   27-AUG-1999;  99US-0151065P.
PR   27-AUG-1999;  99US-0151066P.
PR   27-AUG-1999;  99US-0151080P.
PR   30-AUG-1999;  99US-0151303P.
```

```
PR    31-AUG-1999;    99US-0151438P.
PR    01-SEP-1999;    99US-0151930P.
PR    07-SEP-1999;    99US-0152363P.
PR    10-SEP-1999;    99US-0153070P.
PR    13-SEP-1999;    99US-0153758P.
PR    15-SEP-1999;    99US-0154018P.
PR    16-SEP-1999;    99US-0154039P.
PR    20-SEP-1999;    99US-0154779P.
PR    22-SEP-1999;    99US-0155139P.
PR    23-SEP-1999;    99US-0155486P.
PR    24-SEP-1999;    99US-0155659P.
PR    28-SEP-1999;    99US-0156458P.
PR    29-SEP-1999;    99US-0156596P.
PR    04-OCT-1999;    99US-0157117P.
PR    05-OCT-1999;    99US-0157753P.
PR    06-OCT-1999;    99US-0157865P.
PR    07-OCT-1999;    99US-0158029P.
PR    08-OCT-1999;    99US-0158232P.
PR    12-OCT-1999;    99US-0158369P.
PR    13-OCT-1999;    99US-0159293P.
PR    13-OCT-1999;    99US-0159294P.
PR    13-OCT-1999;    99US-0159295P.
PR    14-OCT-1999;    99US-0159329P.
PR    14-OCT-1999;    99US-0159330P.
PR    14-OCT-1999;    99US-0159331P.
PR    14-OCT-1999;    99US-0159637P.
PR    14-OCT-1999;    99US-0159638P.
PR    18-OCT-1999;    99US-0159584P.
PR    21-OCT-1999;    99US-0160741P.
PR    21-OCT-1999;    99US-0160767P.
PR    21-OCT-1999;    99US-0160768P.
PR    21-OCT-1999;    99US-0160770P.
PR    21-OCT-1999;    99US-0160814P.
PR    21-OCT-1999;    99US-0160815P.
PR    22-OCT-1999;    99US-0160980P.
PR    22-OCT-1999;    99US-0160981P.
PR    22-OCT-1999;    99US-0160989P.
PR    25-OCT-1999;    99US-0161404P.
PR    25-OCT-1999;    99US-0161405P.
PR    25-OCT-1999;    99US-0161406P.
PR    26-OCT-1999;    99US-0161359P.
PR    26-OCT-1999;    99US-0161360P.
PR    26-OCT-1999;    99US-0161361P.
PR    28-OCT-1999;    99US-0161920P.
PR    28-OCT-1999;    99US-0161992P.
PR    28-OCT-1999;    99US-0161993P.
PR    29-OCT-1999;    99US-0162142P.
```

```
Query Match           72.5%;     Score 1913; DB 3;   Length 449;
Best Local Similarity 82.5%;     Pred. No. 2.1e-180;
Matches 363;  Conservative 35;  Mismatches 30;  Indels 12;  Gaps  6;

Qy       68 MLCAGSVRAKNTMNIMLTNVLDAAAGGLFYYLFGFAFGSPSNGFIGKHFFGLKDIPSS 127
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||::
Dna       1 MLCAGSVRAKNTMNIMLTNVLDAAAGGLFYYLFGYAFGSPSNGFIGKHYFGLKDIPTA  60

Qy      128 SYDYSYFLYQWAFAIAAAGITSGSIAERTQFVAYLIYSSFLTGFVYPVVSHWFWSPDGWA 187
            |||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||
Db       61 SADYSNFLYQWAFAIAAAGITSGSIAERTQFVAYLIYSSFLTGFVYPVVSHWFWSVDGWA 120

Qy      188 SAFKIT-DRLFSTGVIDFAGSGVVHMVGGIAGLWGALIEGPRMGRFDHAGRAVALRGHSA 246
            ||:|||  ||||| ||||||||||||||||||||||||||:|||:|| ||||||||||||
Db      121 SPFRTDGDLLFSTGAIDFAGSGVVHMVGGIAGLWGALIEGPRLGRFDNGGRAIALRGHSA 180

Qy      247 SLVVLGTFLLWFGWYGFNPGSFNKILLVTY-ETGTYNGQWSAVGRTAVTTLAGSTRALTT 306
            ||||||||||||||||||||||||||:||:   |||||||||||||||||||||:||||
Db      181 SLVVLGTFLLWFGWYGFNPGSFNKILVTY ETGTYNGQWSAVGRTAVTTTLAGCTRALTT 239

Qy      307 LFGKRVISGHWNVTDVCNGLLGGFAAITAGCSVVEPWAAIVCGFVASIVLIACNKLAEKV 366
            ||||::|||||||||||||||||||||||||:||:|||||||||||| |:|||||||:|:
Db      240 LFGKRLLSGHWNVTDVCNGLLGGFAAITGGCSVVEPWAAIICGFVAALVLLGCNKLAEKL 299

Qy      367 KFDDPLEAAQLHGGCGTWGVIFTALFAKKEYVKEVYG--LGRAHGLLMGGGGKLLAAHVI 424
            |||||||||||||||| |||||||||||||::|||   :|||:||| |||||||||||||
Db      367 RGPLSNPFIAIGDF-LLDPPVGDALMGQIFAAFLFQLSFATTATTIVSGAMAERCNFKAY 234
```

```
Db      300 KYDDPLEAAQLHGGCGAWGLIFTALFAQEKYLNQIYGNKPGRPHGLFMGGGGKLLGAQLI 359
              |:||||||||||||| |||||||||:|||||||||| || |||||||||||||| ||::|
Qy      425 QILVIAGWVSATMGPLFWGLNKLKLLRISSEDELAGMDMTRHGGFAYAY-EDDETHKHGM 483
              ||:|||||||||||| ||||||||||||::|| ||::||||||||||  :||||::|| ::
Db      360 QIIVITGWVSATMGTLFFILKKMKLLRISSEDEMAGMDMTRHGGFAYMYFDDDESHK-AI 418

Qy      484 QLRRVGP-----NASSTPT 497
              |||: ||     :|::||
Db      419 QLRRVEPRSPSPSGANTTPT 438
```

```
RESULT 10
ABB62422
ID   ABB62422  standard; protein; 651 AA.
XX
AC   ABB62422;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Drosophila melanogaster polypeptide SEQ ID NO 14058.
XX
KW   Drosophila; developmental biology; cell signalling; insecticide;
KW   pharmaceutical.
XX
OS   Drosophila melanogaster.
XX
PN   WO200171042-A2.
XX
PD   27-SEP-2001.
XX
PF   23-MAR-2001; 2001WO-US009231.
XX
PR   23-MAR-2000; 2000US-0191637P.
PR   11-JUL-2000; 2000US-0614150.
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Venter JC,  Adams M,  Li PWD,  Myers EW;
XX
DR   WPI; 2001-656860/75.
DR   N-PSDB; ABL06525.
XX
PT   New isolated nucleic acid detection reagent for detecting 1000 or more
PT   genes from Drosophila and for elucidating cell signaling and cell-cell
PT   interactions.
XX
PS   Disclosure; SEQ ID NO 14058; 21pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated nucleic acid detection reagent
CC   capable of detecting 1000 or more genes from Drosophila. The invention is
CC   useful in developmental biology and in elucidating cell signalling and
CC   cell-cell interactions in higher eukaryotes for the development of
CC   insecticides, therapeutics and pharmaceutical drugs. The invention
CC   discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC   sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC   ABB72072). The sequence data for this patent did not form part of the
CC   printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 651 AA;

Query Match           24.8%;     Score 655.5; DB 4;   Length 651;
Best Local Similarity 36.1%;     Pred. No. 1.3e-55;
Matches 155;  Conservative 73;  Mismatches 164;  Indels 37;  Gaps  14;

Qy       48 DNTYLLFSAYLVFSMQLGFAMLCAGSVRAKNTMNIMLTNVLDAAAGLFYYLFGFAFAFG 107
            ||::|:|| :||:|:|::|||||||||||||||||||||||||||||||||||:|:|:||
Db      116 DTNWVLTSSFIIFTMQTGFGMLESGCVSIKNEVNIMKQNVIDIVLGGFTYWLFGYGMSFG 175

Qy      108 -SP-SNGFIGKHFFGLKDIPSS----SYDYSYFLYQWAFAIAAAGITSGSIAERTQFVAY 161
             :|  ||||||||||||:||    :|||::||||||||||||||||||||:||||||||
Db      176 RGPLSNPFIAIGDF-LLDPPVGDALMGQIFAAFLFQLSFATTATTIVSGAMAERCNFKAY 234
```

```
Qy  162 LIYSSFLTGFVYPVVSHWFWSPDGWASAFKITDRLFSTGVIDFAGSGVVHMVGGIAGLWG 221
        ::|||  ||:||  :       :  : :||||| ||::|| :
Db  235 CLF-SFLNTAVVCIPAGWVWGEHGF--------LNKLGAVDIAGSGPVHLIGGASAFAS 284

Qy  222 ALIEGPRMGRFDHAGRAVALRGHSASLVVLGTFLLWFGWYGFENPGSFNKILLTYGNSGNY 281
        :: :|| ||  |:: : ||| :|||||  ||||  ||::| :
Db  285 AAMLGPRLGRYSEGYDPLPL-GNPVN-ACMGLFVLMWGWLAFNSGS------TYGVSG-- 334

Qy  282 YGQWSAVGRTAVTTL----AGSTAALTLFGKRVISGHWNVTDVCNGLLGGFAAITAGC 337
          :     |  :|| |   :         :|| :::|| :|: ||  |:|||||||
Db  335 -AKWQYAARAAVWTMMGSPGGGFTSSIYSFWRH--GGGMDIMDLINGVLGSLVSTTAGC 390

Qy  338 SVVEPWAAIVCGFVASIVLIACNKLAEKVKFPDDPLEAAQLHGGCGTWGVIFTALFAKKEY 397
         : |||| |:|| ::|| :::| :: :|||| ||::|||| :|||:|||||
Db  391 FLYRAWEALVIGAIGSLFCVLAMPLFDRMGVDDPVGASAVHGVCGIWGVIAVGLFADNPI 450

Qy  398 VKEVYGLGRAHGLLMGGGKLLAAHVIQILVIAGWVSATMGPLFWGLNLKLLRISSEDE 457
         |: |||||||||||| |||||||||||||||||||||||||| ||||||:|| ||
Db  451 PLDTTN-GRS-GLPKGGGWYLLGIGTLSALCLACWGVCSTFLLLYVINKIIPIRMDPHEE 508

Qy  458 LAGMDMTRH 466
        | || |||
Db  509 LLGADLTEH 517
```

RESULT 11
AAO26532
ID AAO26532 standard; protein; 183 AA.
XX
AC AAO26532;
XX
DT 18-FEB-2003 (first entry)
XX
DE Corn high affinity ammonium transporter protein.
XX
KW Herbicide; ammonium transporter protein; herbicide; transgenic plant;
KW corn.
XX
OS Zea mays.
XX
PN US2002142390-A1.
XX
PD 03-OCT-2002.
XX
PF 28-DEC-2001; 2001US-00033109.
XX
PR 28-AUG-1998; 98US-0098248P.
PR 27-AUG-1999; 99US-00384625.
XX
PA (ALLE/) ALLEN S M.
PA (RAFA/) RAFALSKI J A.
XX
PI Allen SM, Rafalski JA;
XX
DR WPI; 2003-102520/09.
DR N-PSDB; AAL53984.
XX
PT Novel ammonium transporter polypeptide useful for identifying enzymatic
PT inhibitors, which is homologous to corn, soybean, wheat or rice ammonium
PT transporter polypeptides.
XX
PS Claim 14; Page 13; 27pp; English.
XX
CC The invention relates to a novel ammonium transporter protein comprising
CC 90% homology based on the Clustal method compared to: a corn ammonium
CC transporter polypeptide of 183 or 63 amino acids; a soybean ammonium
CC transporter polypeptide of 500 or 486 amino acids; a wheat ammonium
CC transporter polypeptide of 494 or 470 amino acids; and a rice ammonium
CC transporter 497 amino acids fully defined in the specification. The
CC isolated polynucleotide is useful for selecting an isolated
CC polynucleotide that affects the level of expression of the ammonium
CC transporter polypeptide in a plant cell. The ammonium transporter protein
CC is useful for preparing antibodies which are useful for detecting the
CC transporter protein in situ in cells or in vitro in cell extracts, and as
CC targets to facilitate design and/or identify inhibitors of the enzymes
CC that are useful as herbicides. The isolated polynucleotide is also useful
CC for creating transgenic plants in which the polynucleotide is present at
CC higher or lower levels than normal, and for designing and producing
CC primer pairs which are useful in amplification or primer extension
CC reactions. This sequence represents the corn high affinity ammonium
CC transporter protein of the invention
XX
SQ Sequence 183 AA;

Query Match   23.2%; Score 613; DB 6; Length 183;
Best Local Similarity   65.5%; Pred. No. 3.6e-52;
Matches 116; Conservative 23; Mismatches 30; Indels 8; Gaps 2;

```
Qy  330 FAAITAGCSVVEPWAAIVCGFVASIVLIACNKLAEKVKFPDDPLEAAQLHGGCGTWGVIFT 389
        |||||||||||||||||||||||||||||||||||||||||||||||||||:||||||||
Db    3 FAAITAGCSVVEPWAAVICGFVSAWVLIGANALAARFRFDDPLEAAQLHGGCGAWGVLFT 62

Qy  390 ALFAKKEYVKEVYGLGRAHGLLMGGGKLLAAHVIQILVIAGWVSATMGPLFWGLNKLKL 449
        ||::||:|||| :|| |:|||||||||||||||||||||||||||||||||||||||:|
Db   63 GLFARRKVEEIYGAGRPYGLFMGGGKLLAAQIIQILVIAGWVSCTMGPLFYALKKUGL 122

Qy  450 LRISSEDELAGMDMTRHGGFAYAYED----DETHKHGMQLR----RVGPNASSTPTT 498
        ||||::|||:|||||||||||:|:|         ::|  |    :| |: || ||||
Db  123 LRISADDEMSGMDLTRHGGFAYVYHDEDPGDKAGVGGFMLKSAQNRVEPAAAVAAAT 179
```

RESULT 12
ABB49278
ID ABB49278 standard; protein; 401 AA.
XX
AC ABB49278;
XX
DT 05-FEB-2002 (first entry)
XX
DE Listeria monocytogenes protein #1982.
XX
KW Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
KW vitamin B12; bacterial infection; disease.
XX
OS Listeria monocytogenes.
XX
PN WO200177335-A2.
XX
PD 18-OCT-2001.
XX
PF 11-APR-2001; 2001WO-FR001118.
XX
PR 11-APR-2000; 2000FR-00004629.
XX
PA (INSP) INST PASTEUR.
XX
PI Buchrieser C, Frangeul L, Couve E, Rusniok C, Fsihi H, Dehoux P;
PI Dussurget O, Chetouani F, Nedjari H, Glaser P, Kunst F, Cossart P;
PI Daniels J, Goebel W, Kreft J, Kuhn M, Ng E, Vazquez-Boland JA;
PI Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A;
PI Chakraborty T, Domann E, Hain T, Berche P, Charbit A, Durant L;
PI Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N;
PI Maduenio B, De Pablos B, Wehland J, Kaerst U, Entian K, Hauf J;
PI Rose M, Voss H;
XX
DR WPI; 2002-010914/01.
XX
PT Genomic sequence for Listeria monocytogenes, useful e.g. for treatment
PT and prevention of Listeria and related bacterial infections, and related
PT polypeptides.
XX
PS Claim 6; SEQ ID NO 1983; 192pp; French.
XX
CC The present invention relates to the genome sequence of Listeria
CC monocytogenes EGD-e (see ABA03041). The genome sequence and fragments of

```
CC    it are useful for selecting probes and primers for detecting genes in L.
CC    monocytogenes and related organisms, and for studying genetic
CC    polymorphisms and other genomes. The present sequence is a protein
CC    encoded by the genome sequence of the present invention. Proteins
CC    expressed from the genome sequence are useful for raising specific
CC    antibodies, identification of L. monocytogenes and related organisms, and
CC    for biosynthesis and biodegradation, especially biosynthesis of Vitamin
CC    B12. The genome sequence and proteins encoded by it are also useful for
CC    selecting compounds that regulate gene expression and cell replication
CC    and modulate L. monocytogenes-related diseases. In addition, the genome
CC    sequence and proteins encoded by it are useful in pharmaceutical and
CC    vaccines compositions for the treatment or prevention of infections by L.
CC    monocytogenes and related organisms. Note: The sequence data for this
CC    patent did not form part of the printed specification, but was obtained
CC    in electronic format directly from WIPO at
CC    ftp.wipo.int/pub/published_pct_sequences
XX
SQ    Sequence 401 AA;

Query Match            18.6%;   Score 491;  DB 5;  Length 401;
Best Local Similarity  31.7%;   Pred. No. 1.3e-39;
Matches  142;  Conservative   66;  Mismatches  168;  Indels   72;  Gaps   17;

Qy     47  VDNTYLLFSAYLVFSMQLGFAMLCAGSVRAKNTMNIMLTNVLDAAAGGLFYYLFGFAFAF  106
               ::::|:||:|  |:||||:| ||:: : :  :  : :  :|:  |::|::||  :|: ||
Db      1  MESVFMFFCTLLVWLMTPGIALFYGGMVRRKNVLSTAMYSFSSMAIISLLWIVGYSLAF   60

Qy    107  GSPSNGFIGK------HFFGLKDIPSSSYDYS------YFLYQWAFAIAAAGITSGSGSIA  153
               :|:|  ||::          ||: |  : ::        :|:|  | ::| ::|:: |
Db     61  -APGNGFIGSPDWTFLHNVGF---AANATYSDAIPHVLFMMFQMTFAILTVAIISGAFA  115

Qy    154  ERTQFVAYLIYSSFLTGFVYPVVSHWFWSPDGWASAFKITDRLFSTGVIDFAGSGVVHMV  213
               ||||||  : |:|||| |  |||   ||   |  |  ||    || ||:|||||||||:
Db    116  ERMNFSAYLIFILWSLLVYSPVAHWVWGDGGW---------LRRNLGALIDFAGGNVVHIS  166

Qy    214  GGIAGLWGALIEGPRMGRFDHAGRAVALRGHSASLVVLGTFLLWFGWYGFNPGSFNKILL  273
               |||| |||:|| | ::|||   |  : |||| ::||:|| |||||| ||| :|: |:|
Db    167  SGVTGLVLAIMGRRK------EADSASPHNLPLALIGGILVWFGWYGFNVGS----AL  215

Qy    274  TYGNSGNYYGQWSAVGRTAVTTLAGSTAALTTLFGKRVISGHW------NVTDVCNGLL  327
               | |  |||||   :|| |||||||| | ||| ||||||||||       ||||||||:
Db    216  TIDN--------VAMTAFVNT--NTAAAGIIGWGLV--EWLTNKKPTMLGTISGAI     260

Qy    328  GGFAAITAGCSVVEPWAAIVCGFV-ASIVLIACNKLAEKVKFDDPLEAAQLHGGCGTWGV  386
               ||||||||| :  :|:|   | || |: :|  | :| :: || || : | |||| : | |
Db    261  AGLVSITPAAGFVTVPSSLIIGFFGGALCFWAVFWLKGKVKYDDALDAFGLHGIIGWGG  320

Qy    387  IFTALFAKKEYVKEVYGLGRAHGLLMGGGGKLLAAHVIQILVIAGWVS--ATMGPLFWGL  444
               ||| || || ||| ||||| ::| ||||| |:|| :| :::||  :    |:|||||| |
Db    321  IATGLFATTK-VNE---AGADGLFYGNASLIV----VKQLIAIGSTVAYVAVTALIVVI  371

Qy    445  NKLKL-LRISSEDELAGMDMTRHGGPAY  471
               :|||  || |:||:|  :|  |:|  ||
Db    372  IKLFLPIRVNEEQEYKGLDLTLHGEKAY  399
```

RESULT 13
ADA34110
```
ID   ADA34110 standard; protein; 470 AA.
XX
AC   ADA34110;
XX
DT   20-NOV-2003  (first entry)
XX
DE   Acinetobacter baumannii protein #1271.
XX
KW   Acinetobacter baumannii; bacterial disease; antibacterial; vaccine;
KW   plant biocontrol agent.
OS   Acinetobacter baumannii.
XX
PN   US6562958-B1.
XX
```

```
PD   13-MAY-2003.
XX
PF   04-JUN-1999;   99US-00328352.
XX
PR   09-JUN-1998;   98US-0088701P.
XX
PA   (GENO-) GENOME THERAPEUTICS CORP.
XX
PI   Breton G, Bush D;
XX
DR   WPI; 2003-576092/54.
DR   N-PSDB; ADA29984.
XX
PT   New Acinetobacter baumanii proteins and nucleic acids, useful as reagents
PT   for diagnosing a bacterial disease, as components of antibacterial
PT   vaccines, as targets for antibacterial drugs, or as biocontrol agents for
PT   plants.
XX
PS   Example; SEQ ID NO 5397; 328pp; English.
XX
CC   The invention relates to isolated Acinetobacter baumannii nucleic acids.
CC   The A. baumannii nucleic acids and polypeptides are useful as reagents
CC   for diagnosing a bacterial disease, as components of antibacterial
CC   vaccines, as targets for antibacterial drugs, to detect the presence of
CC   A. baumannii and other Acinetobacter species in a sample, in screening
CC   compounds for the ability to interfere with the A. baumannii life cycle
CC   or to inhibit A. baumannii infection, and as biocontrol agents for
CC   plants. The present sequence represents the amino acid sequence of an A.
CC   baumannii protein.
XX
SQ   Sequence 470 AA;

Query Match            18.5%;   Score 488;  DB 6;  Length 470;
Best Local Similarity  30.0%;   Pred. No. 3.3e-39;
Matches  144;  Conservative   71;  Mismatches  179;  Indels   86;  Gaps   16;

Qy     20  TDASAAASLICGHFAAVDSKFVDTAFAV---DNTYLLFSAYLIVFSMQL-GFAMLCAGSV   74
               |||:  :| | |::| ||  |  |||    ||| || ||: |:|:| ||||:||||:|
Db     41  TSTSEAPATIAAAPAAEETPAAPTPTAKLDTGDTSWILISTALVLLMTIPGLALFYGGMV  100

Qy     75  RAKNTMNIMLTNVLDAAAGGLFYYLFGFAFAFGSPSNGFIGKHFFGLK--------DIP  125
               | || |: :   |  :|| |||| ||| |:     : :    ||    :||
Db    101  RKKGNVLSTMMFSLSAAILVSLLWVIAGYSIAFSG-:-----TGAYFGDLSKAMLNGVAFDAL  155

Qy    126  SSSYDYSYF-LYQWAFAIAAAGITSGSIAERTQFVAYLIYSSFLTGFVYPVVSHWFWSPD  184
               |: :  :|   | :|  |  || ::||    || |||:|:|:  |||| :|:|  ||
Db    156  SGTIPESLFVIFQMTFAIITVALLSGSIADRMKYSAFMAFIAIWVLVVYAPITHWTWAAD  215

Qy    185  GWASAFKITDRLFSTGVIDFAGSGVVHMVGGIAGLWGALIEGPRMGRFDHAGRAVALRGH  244
               ||   ::|:   |::|  :||||:||:| |::   |  |::|| |||||||| ::|  |:
Db    216  GW--------LFKAGALDFAGGTVVHINSGVAGLVAAYMLGKRIG---LGRE-SMAPH  261

Qy    245  SASLVVLGTFLIWFGWYGFNPGSFNKILLTYGNSGNYYGQWSAVGRTAVTTLAGSTAAL  304
               :|:|::| |  :|:|:|||| ||   |: :|| |  :  |:     |  |:|||:|| |
Db    262  NLTLTVIGASLLWVGWFGFNGGS------------------ALGAGARASMAILVTQVAAAAAAF  308

Qy    305  TTLFGKRVISGHWNVTDVCNGLLGGFAAITAGCSVVEPWAAIVCGFVASIVLI-ACNKLA  363
               : ::|| || :: :|::|: | ||:|    |:  :|     |:   |  :|: | ::|
Db    309  SWLVVERMIRGKASVLGGASGAVAGLVVITPAAGFVGVGGALVMGLIGGVVCFWGITALK  368

Qy    364  EKVKFDDPLEAAQLHGGCGTWGVIFTALFAKKEYVKEVYGLGRAHGLLMGGGGKLLAAHV  423
               |::|    :: |:  ||| |: |  ||| ||||| |   |::| |||||||| |::|  |
Db    369  RLLKADDALDAFGLHAVGGIVGAILTGVFYSDEIIK--------------------AANV  408

Qy    424  IQILVIAG--WVS-----ATMGPLFWGLNK--LKL------LRISSEDELAGMDMTRHG  467
               | |:|::| |:|     ::|:|||| |:|  || :       || |:|:   :|     |
Db    409  ALAPTFAGQLWVQVEGVLATM--VYSGIATFIILKVIDLIIGLRVNSDDERMGLDLSQHG  466
```

RESULT 14
ABP65520
```
ID   ABP65520 standard; protein; 431 AA.
XX
```

```
AC    ABP65520;
XX
DT    19-NOV-2002  (first entry)
XX
DE    Bifidobacterium longum NCC2705 ORF amino acid sequence SEQ ID NO:264.
XX
KW    Bifidobacterium longum NCC2705; Bifidobacterium; bacterial;
KW    antidiarrheic; antibacterial; inhibitor of Salmonella; detection;
KW    identification; lactic acid bacterium; diarrhoea; pathogenic bacteria;
KW    rotavirus; food composition; pharmaceutical composition.
XX
OS    Bifidobacterium longum.
XX
PN    EP1227152-A1.
XX
PD    31-JUL-2002.
XX
PF    30-JAN-2001; 2001EP-00102050.
XX
PR    30-JAN-2001; 2001EP-00102050.
XX
PA    (NEST ) SOC PROD NESTLE SA.
XX
DR    WPI; 2002-668397/72.
XX
PT    Novel polynucleotide comprising Bifidobacterium genome sequence useful as
PT    a probe or primer for detecting and/or identifying Bifidobacterium longum
PT    in a biological sample.
XX
PS    Claim 3; SEQ ID NO 264; 80pp; English.
XX
CC    The present invention describes a polynucleotide (I) comprising a
CC    sequence of a Bifidobacterium genome selected from the nucleotide
CC    sequences given in ABQ81842 and ABQ81843, or a sequence exhibiting at
CC    least 90% identity or which hybridises with the sequences given in
CC    ABQ81842 and ABQ81843. Also described is a polynucleotide (II) encoding a
CC    fusion protein, comprising a sequence selected from 1097 sequences given
CC    in ABP65258 to ABP6354 ligated in frame to a polynucleotide encoding a
CC    heterologous polypeptide. (I) has antidiarrheic and antibacterial
CC    activities, and can be used as an inhibitor of Salmonella. (I) (which is
CC    a probe) is useful for the detection and/or identification of
CC    Bifidobacterium longum in a biological sample. A carrier containing the
CC    lactic acid bacterium Bifidobacterium longum NCC2705 (CNCM I-2618) can be
CC    used for preventing and/or treating diarrhoea brought about by pathogenic
CC    bacteria and/or rotavirus. The carrier is a food composition selected
CC    from milk, yogurt, curd, cheese, fermented milks, milk based fermented
CC    products, ice-creams, fermented cereal based products, milk based
CC    powders, infant formula, pet food or a pharmaceutical composition
CC    selected from tablets, liquid bacterial suspensions, dried oral
CC    supplement, wet oral supplement, dry tube feeding or wet tube feeding.
CC    (I) is useful in DNA arrays or chips to carry out analysis of the
CC    expression of the Bifidobacterium gene. ABQ81844 to ABQ81850 represent
CC    Bifidobacterium related nucleotide sequences given in the Sequence
CC    Listing from the present invention but not mentioned further within the
CC    specification. N.B. The sequence data for this patent is not represented
CC    in the printed specification but is based on sequence information
CC    supplied by the European Patent Office
XX
SQ    Sequence 431 AA;

Query Match             18.2%;  Score 481.5;  DB 5;  Length 431;
Best Local Similarity   29.8%;  Pred. No. 1.3e-38;
Matches 134;  Conservative  72;  Mismatches 190;  Indels  53;  Gaps  11;

Qy     46 AVDN--TYLLFSAYLLVFSMQLGFAMLCAGSVRAKNTMNILMLTNVLDAAAGGLFYYLFGF 102
            |:|:   |||||  |||||| |||||||      :|:|:|:    . :|: | |:|:
Db      3 AIDSGNTAWILTSASLVFLMTPGVAFFYGGMVRAKAVLNMLMLEAAALSVTMIVWTLWGW 62

Qy    103 APA--------PGSPSNGFIGKHFFGLKD------IPSSSYDYSY-FLYQWAFAIAAA 145
            ||:        |:||:|  .:|:| |||      :|  :| ||| |||| ||||:
Db     63 SIAYAGSDIGGIFGDPASGFLLKDSMVAQDGVFTATGLNGNNYPVSVDVAFQVAFAMITV 122

Qy    146 GITSGSIAERTQFVAYLLIYSSFLITGFVYPVSHWFWSPDGWASAFKITDRLFSTGVIDFA 205
```

```
Db    123 GLICGAIAERVKYSTWMIFVALMVTFDYAPMAHMVWN-GGLLSADGAISKAIGAAAHDFA 181
            |: |:||||  ::|:.   ::|:.    |  ||    ||:.
Qy    206 GSGVVHMVGGIAGLWGALIEGPRMGRFDHAGRAVALRGHSASLVVLGTFLLWFGYGFNP 265
            |||:  ::|:        ||:::||           :|||| ||::|||||||:|||:||
Db    182 GGTVVHINAAVAALIVLIIGRKG-----FGTQPFRPHNVPFVMLGAFLLWFGWFGFNA 236
            ||||::|     :|||: :||::           |   |::  :|||||||::|||||
Qy    266 GSFNKILLTYGNSGNYYGQWSAVGRTAVTTLAGSTAALTTLFGKRVISGHWNVTDVCNG 325
            ||:  ||||:|||:|||: ||||| |||:::| |:::: |||||:| :|||||||  :|
Db    237 GS-----AFAANGTAGYAW------VSTSAATAAAMLSWGFTEKIRSGHYTAMGAASG 283
            ||     | ||| | |:|      |:|:::| |  ||:| ||| :||:|| |||:||
Qy    326 LLGGFAAITAGCSVVEPWAAIVCGFVASIV-LIACNKLAEKVKFPDDPLBAAQLHGGCGTW 384
            ||||| :||| :| ||:| :: |||||:|  ||::|:   :::| |    ||  |||:|
Db    284 MVAGLVAITPAADVVSPLWAIVMGAIAGVLTCLACG-LKFKFGYDDSLDVVGVHGVGGFT 342
            ::|||| || | |||:|| ||:| |:::|| | ||   ::|:|||:|:|:||:|||:||
Qy    385 GVIFTALFAKKEYVKEVYGLGRAHGLIMGGGGKLLAAHVIQIIVIAGWVSATMGPLFWGL 444
            |::|| || :|| : |:|:|| | ||:||||||:||||:|:|:|:  ::|::| |||:||
Db    343 GTILIGFF-----------GGGTGLLAGGDWKQLVVQLVVIALVAILYSAVITAIIAFAL 390
            || |:||           ||| ||||| |||:|:|:|:|:||||: |||::||:||:||
Qy    445 NKLKLLRISSEDELAGMDMTRHGGFAYAY 473
            ::|: ||:::|  |:|  :  |||:||:|
Db    391 EKTIGWRVTEAQEVGGIDLADQGBRAYDF 419
```

```
RESULT 15
ABB54938
ID   ABB54938 standard; protein; 413 AA.
XX
AC   ABB54938;
XX
DT   29-AUG-2003  (revised)
DT   16-MAY-2002  (first entry)
XX
DE   Lactococcus lactis protein amtB.
XX
KW   Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese.
XX
OS   Lactococcus lactis; IL1403.
XX
PN   FR2807446-A1.
XX
PD   12-OCT-2001.
XX
PF   11-APR-2000; 2000FR-00004630.
XX
PR   11-APR-2000; 2000FR-00004630.
XX
PA   (INRG ) INRA INST NAT RECH AGRONOMIQUE.
XX
PI   Bolotine A, Sorokine A, Renault P, Ehrlich SD;
XX
DR   WPI; 2002-043418/06.
XX
PT   New nucleotide sequence useful in the identification or Lactococcus
PT   lactis and related species.
XX
PS   Claim 6; SEQ ID NO 1640; 2504pp; French.
XX
CC   The present invention is related to a Lactococcus lactis nucleotide
CC   sequence (ABA90521) and related proteins (ABB53300-ABB55621). The nucleic
CC   acid sequence is useful in the detection and/or amplification of nucleic
CC   acid sequence, particularly to identify Lactococcus lactis or related
CC   species. The proteins of the invention are useful for the biosynthesis or
CC   biodegradation of a composition of interest. The invention helps research
CC   in lactic bacteria, particularly useful in the production of yogurt and
CC   cheese. Note: The sequence data for this patent is based on equivalent
CC   patent WO200177334 (published 18-OCT-2001) which is available in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences. (Updated on 29-AUG-2003 to
CC   standardise OS field)
XX
SQ   Sequence 413 AA;
```

Wed Mar 10 09:05:08 2004

us-10-033-109-4.rag

Query Match   17.1%;   Score 452.5;   DB 5;   Length 413;
Best Local Similarity   29.3%;   Pred. No. 9e-36;
Matches 129;   Conservative   73;   Mismatches 177;   Indels   61;   Gaps   14;

Qy   51 YLLFSAYLVFSMQLGFAMLCAGSVRAKQTMNIMLTNVLDAAAGGLFYYLFGFAFAFGSPS 110
            ::|:|||| | | : | |:::: | : :|||:::|:|:| |
Db   8 FILICAALVFLMTPALAFFYGGLGRRKNVLNTMMMSLAPMALASILWIIIGFSFSF-SGS 66

Qy   111 NGFIG--KHFF--GLKDIPSSSY------DYSYFLYQWAFAIAAAGITSGSIAERTQFVA 160
           |::|: | | | :   ::   :| | : :| : : :
Db   67 NSWIGDFHHLFMNGVDMAKNSLFPANHIPDGLFSGFQMMFSIITVALITGSVVGRMRFTP 126

Qy   161 YLIYSSFLTGFVYPVVSHWFWSPDGWASAFKITDRLFSTGVIDPAGSGVVHMVGGIAGLW 220
           |||: | : :| ||||||| ||| :| |
Db   127 ILIFMTAWLIVYYPLAHMV-----WGGGF------LAQIHA-IDFAGGDVVHISSGVTGLG 176

Qy   221 GALIEGPRMG--RFDHAGRAVALRGHSASLVVLGTFLLWFGWYGFNPGSFNKILLTYGNS 278
           ||:| :| : :| |||||:||:||:||| : :
Db   177 LALVLGKRRDYERLDY-------RPHNIPFVVLGAGLLWFGFNAGS---ALAANGVA 226

Qy   279 GNYYGQWSAVGRTAVTTLAGSTAALTTLFGKRVISGHWNVTDVCNGLLGGFAAITAGCS 338
           :|| :: ||| :| |:::   : |: |:
Db   227 INAF--------MTTNTAAAAAMFSWMIVEKILIGKPSIVGACSGAVVGLVAITPGAG 276

Qy   339 VVEPWAAIVCGFVAS-IVLIACNKLAEKVKFDDPLEAAQLHGGCGTWGVIFTALFAKKEY 397
           |::|:|: | | :| :|::| | :|:|| : | | | |: :|||
Db   277 FVSLWSSLIIGLLVSPLSYFMISVVKKLGYDDALDAFGCHGIGGMFGGIMTGIFATPAL 336

Qy   398 VKEVVGLGRAHGLLMGGGGKLLAAHVIQIL-----VIAGWVSATMGPLFWGLNKLKLLR 451
           : :::|:|:|||::   : :| : :: ::| |:: | |
Db   337 APE------KGYAGLIYGSGKLLLANVSAVVFTVIFTALVSWIIIKVIALFMP------IR 385

Qy   452 ISSEDELAGMDMTRHGGFAY 471
           :: :|:||
Db   386 VSDRAEAIGLDDSEHEETAY 405

Search completed: March 9, 2004, 11:54:06
Job time : 64 secs

```
                GenCore version 5.1.6
        Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on:    March 9, 2004, 11:52:48 ; Search time 21 Seconds
                                     (without alignments)
                                     2290.274 Million cell updates/sec

Title:      US-10-033-109-4
Perfect score: 2639
Sequence:   1 MSLPACPAEQLAQLLGPNTT.........HGMQLRRVGPNASSTPTTDE 500

Scoring table: BLOSUM62
               Gapop 10.0 , Gapext 0.5

Searched:   283366 seqs, 9619152G residues

Total number of hits satisfying chosen parameters:    283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :  PIR_78:*
            1: Pir1:*
            2: Pir2:*
            3: Pir3:*
            4: Pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2129.5	80.7	501	2	T06653	ammonium transport
2	1973.5	74.8	504	2	T04529	probable ammonium
3	1953	74.0	514	2	T06585	ammonium transport
4	1919	72.7	514	2	A96671	Ammonium transport
5	1792.5	67.9	533	2	T03441	probable ammonium
6	771.5	29.2	435	2	H72379	ammonium transport
7	715.5	27.1	498	2	AC1930	ammonium transport
8	711.5	27.0	468	2	A69468	ammonium transport
9	706	26.8	534	2	T15414	hypothetical prote
10	703	26.6	507	2	S76687	hypothetical prote
11	678.5	25.7	442	2	S74801	ammonium transport
12	665	25.2	518	2	AD1930	ammonium transport
13	655	24.8	421	2	B84129	ammonium transport
14	579.5	22.0	622	2	T23804	hypothetical prote
15	569.5	21.6	541	2	S76017	hypothetical prote
16	523	19.8	439	2	B75487	ammonium transport
17	523	19.8	470	2	AE1930	ammonium transport
18	523	19.8	592	2	T15413	hypothetical prote
19	504.5	19.1	388	2	D69468	ammonium transport
20	503	19.1	448	2	T35667	ammonium transport
21	500	18.9	405	2	H96983	ammonium transport
22	496	18.8	477	2	H70747	probable ammonium
23	494	18.7	404	2	A36865	ammonium transport
24	492	18.6	401	2	AF1626	ammonium transport
25	491	18.6	401	2	AD1264	ammonium transport
26	485	18.4	420	2	F64467	ammonium transport
27	471.5	17.9	449	2	AE2915	ammonium transport
28	471.5	17.9	449	2	H97689	amtB protein (AJ00
29	470	17.8	391	2	A69372	ammonium transport
30	470	17.8	428	2	C64775	ammonium transport
31	470	17.8	428	2	A90692	probable ammonium
32	470	17.8	428	2	E85542	probable ammonium
33	463.5	17.6	433	2	AB3273	ammonium transport
34	462	17.5	492	2	A82632	ammonium transport
35	460	17.4	407	2	A69188	ammonium transport
36	457	17.3	423	2	H70310	ammonium transport
37	457	17.3	428	2	AC0560	probable ammonium
38	456.5	17.3	412	2	C69188	ammonium transport
39	454.5	17.3	554	2	T22433	hypothetical prote
40	452.5	17.1	413	2	H86823	ammonium transport
41	449.5	17.0	492	2	S46225	ammonium transport
42	446	16.9	489	2	S69027	ammonium transport
43	442	16.7	442	2	C82985	ammonium transport
44	438.5	16.6	431	2	AF0381	probable ammonium
45	436	16.5	391	2	B64307	ammonium transport

ALIGNMENTS

```
RESULT 1
T06653
ammonium transport protein amt1 - Arabidopsis thaliana
N;Alternate names: NH4+ transporter; protein T6G15.60
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 23-Apr-1999 #sequence revision 23-Apr-1999 #text_change 24-Nov-1999
C;Accession: T06653; S46226; S59843
R;Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X
submitted to the Protein Sequence Database, April 1999
A;Reference number: Z15791
A;Accession: T06653
A;Molecule type: DNA
A;Residues: 1-501 <BEV>
A;Cross-references: EMBL:AL049656; GSPDB:GN00062; ATSP:T6G15.60
A;Experimental source: cultivar Columbia; BAC clone T6G15
R;Ninnemann, O.; Jauniaux, J.C.; Frommer, W.B.
EMBO J. 13, 3464-3471, 1994
A;Title: Identification of a high affinity NH(4)(+) transporter from plants.
A;Reference number: S46226; MUID:94341258; PMID:8062823
A;Accession: S46226
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-82, 'EP',85-501 <NIN>
A;Cross-references: EMBL:X75879
R;Frommer, W.B.
submitted to the EMBL Data Library, January 1994
A;Reference number: S59843
A;Accession: S59843
A;Molecule type: mRNA
A;Residues: 1-501 <FRO>
A;Cross-references: EMBL:X75879; NID:g551218; PIDN:CAA53473.1; PID:g551219
C;Genetics:
A;Gene: ATSP:T6G15.60; amt1
A;Map position: 4
C;Superfamily: ammonium transport protein amt1
C;Keywords: ammonium transport; transmembrane protein

Query Match         80.7%;  Score 2129.5; DB 2;  Length 501;
Best Local Similarity  80.9%;  Pred. No. 2.9e-154;
Matches 407; Conservative 42; Mismatches 39; Indels 15; Gaps 7;

Qy      5 ACPAEQLAQLLGPNTTDASAAASLICGHFAAVDSKFVDTAFAVDNTYLLFSAYLVFSMQL 64
            :: |||:|||||||   :::||  ||::|||    | :::|||||||||||||||||||
Db      2 SCSATDLAVLLGPN--ATAAANYICGQLGDVNNKFIDTAFAIDNTYLLFSAYLVFSMQL 58

Qy     65 GFAMLCAGSVRAKNTMNIMLTNVLDAAAGGLFYYLFGFAFAFGSPSNGFIGKHFPGLKDI 124
          ||||||||||||||||||||||||||||||||||||| |||||||||||||||| ||||
Db     59 GFAMLCAGSVRAKNTMNIMLTNVLDAAAGGLFYYLFGYAFAFGSPSNGFIGKHYFGLKDI 118

Qy    125 PSSSYDYSYFLYQWAFAIAAAGITSGSIAERTQFVAYLLYSSFLTGFVPVVSHWFWSPD 184
          |::||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    119 PTASADYSNFLYQWAFAIAAAGITSGSIAERTQFVAYLLYSSFLTGFVPVVSHWFWSVD 178
```

```
Qy  185 GWASAFKIT-DRLFSTGVIDFAGSGVVHMVGGIAGLWGALIEGPRMGRFDHAGRAVALRG 243
            ||||:   :| :||||||:|||||||||||||||||||||||||:|||||: :|||| |||
Db  179 GWASPFRTDGDLLFSTGAIDFAGSGVVHMVGGIAGLWGALIEGPRLGRFDNGGRAILRG 238

Qy  244 HSASLVVLGTFLLWFGWYGFNPGSFNKILLTYGNSGNYYGQWSAVGRTAVTTTLAGSTAA 303
            |||||||||||||||||||||||||||||:|::   ::|| ||||||||||||||||:|||
Db  239 HSASLVVLGTFLLWFGWYGFNPGSFNKILVFY-ETGTYNGQWSAVGRTAVTTTLAGCTAA 297

Qy  304 LTTLFGKRVISGHWNVTDVCNGLLGGFAAITAGCSVVEPWAAIVCGVASIVLIACNKLA 363
            ||||||||::|||||||||||||||||||||:||||||||||||:||::|| ||:|||||
Db  298 LTTLFGKRLLSGHWNVTDVCNGLLGGFAAITGGCSVVEPWAAIICGFVAALVLLGCNKLA 357

Qy  364 EKVKFDDPLEAAQLHGGCGTWGVIFTALFAKKEYVKEVVG--LGRAHGLLMGGGGKLLAA 421
            ||:|:|||||||||||||||:||:||||||::||:| | :    |||:||:|||||||| |
Db  358 EKLKYDDPLEAAQLHGGCGAWGLIFTALFAQEKYLNQIYGNKPGRPHGLFMGGGGKLLGA 417

Qy  422 HVIQILVIAGWVSATMGPLFWGLNKLKLLRISSEDELAGMDMTRHGGFAYAY-EDDETHK 480
            :::||:|||:||||||| ||| |::::|||||||||||:|||||||||||||:|||||:|:
Db  418 QLIQIIVITGWVSATWGTLFFILKKMKLLRISSEDEMAGMDMTRHGGFAYMYFDDDESHK 477

Qy  481 HGMQLRRVGP------NASSTPT 497
            :::||||| |       : |:|||
Db  478 -AIQLRRVEPRSPSPSGANTTPT 499
```

RESULT 2
T04529
probable ammonium transport protein F16A16.190 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 26-Aug-1999
C;Accession: T04529
R;Bevan, M.; Brandt, P.; Dose, S.; Jarke, D.; Scharfe, M.; Schon, O.; Hoheisel, J.; Mewe
submitted to the Protein Sequence Database, February 1999
A;Reference number: Z15376
A;Accession: T04529
A;Molecule type: DNA
A;Residues: 1-504 <BEV>
A;Cross-references: EMBL:AL035353
A;Experimental source: cultivar Columbia; BAC clone F16A16
C;Genetics:
A;Map position: 4
A;Note: F16A16.190
C;Superfamily: ammonium transport protein amt1
C;Keywords: ammonium transport; transmembrane protein

```
Query Match      74.8%;  Score 1973.5;  DB 2;  Length 504;
Best Local Similarity  75.3%;  Pred. No. 2.1e-142;
Matches 374;  Conservative 45;  Mismatches 67;  Indels 11;  Gaps 5;

Qy    5 ACPAEQLAQLL--GPNTTDASAAASLICGHFAAVDSKFVDTAFAVDNTYLLFSAYLVFSM 62
            :|   :||:|    | ||| |: ||:  | ||:|  :|  | || :||||||||||||||:|
Db    6 SCSASDLIPLLSGGANATARAAAEYICGRFDTVAGKFTDAAYAIDNTYLLFSAYLVFAM 65

Qy   63 QLGFAMLCAGSVRAKNTMNIMLTNVLDAAAGGLFYYLFGFAFAFGSPSNGFIGKHFFGLK 122
            |||||||||||||||||||||||||:||||||||||||||||||||||||||||||||: 
Db   66 QLGFAMLCAGSVRAKNTMNIMLTNVIDAAAGGLFYYLFGFAFAFGSPSNGFIGKHFFGMY 125

Qy  123 DIPSSSYDYSYFLYQWAFAIAAAGITSGSIAERTQFVAYLIYSSFLTGFVYPVVSHWFWS 182
            | |   :||  ||||||:|||| |||||||||||||||||||||||||||:|||:||||||
Db  126 DFPQPTFDYPYFLYQWTFAIAAGITSGSIAERTQFVAYLIYSSFLTGLVYPIVSHWFWS 185

Qy  183 PDGWASAFKITDRLFSTGVIDFAGSGVVHMVGGIAGLWGALIEGPRMGRFDHAGRAVALR 242
             ||||| :     ||:||||||||||||||||||||||||||||||:||||:  : |:|||
Db  186 SDGWASPARSENLLFQSGVIDFAGSGVVHMVGGIAGLWGALIEGPRIGRFGVGGKPVTLR 245

Qy  243 GHSASLVVLGTFLLWFGWYGFNPGSFNKILLTYGNS--GNYYGQWSAVGRTAVTTTLAGS 300
            ||||:||||||||||||||||||||||:|:|:|| :    : |||||||||||||||||| 
Db  246 GHSATLVVLGTFLLWFGWYGFNPGSFATIFKAYGETPGSSFYGQWSAVGRTAVTTTLAGC 305

Qy  301 TAALTTLFGKRVISGHWNVTDVCNGLLGGFAAITAGCSVVEPWAAIVCGFVASIVLIACN 360
            |||||||||||::| :|||||||||||||||||||:|||||||||:||||||::|| ||| 
Db  306 TAALTTLFGKRLIDGYWNVTDVCNGLLGGFAAITSGCSVVEPWAALVCGFVAAWLMGCN 365
```

```
Qy  361 KLAEKVKFDDPLEAAQLHGGCGTWGVIFTALFAKKEYVKEVVG--LGRAHGLLMGGGGKL 418
            :||||::||||||||||||||||:|||::| || |::|||||||:  :|: ||:|||
Db  366 RLAEKLQFDDPLEAAQLHGGCGAWGIIFTGLFAEKRYIAEIFGGDPNRPFGLLMGGGGRL 425

Qy  419 LAAHVIQILVIAGWVSATMGPLFWGLNKLKLLRISSEDELAGMDMTRHGGFAYAYEDDET 478
            ||||| ||||||||||||||||||||||||||||:||||||::||||:: :||||||:|:
Db  426 LAAHVQILVITGWVSVTMGTLFFILHKLKLLRIPABDEIAGVDPTSHGGLAYMYTEDEI 485

Qy  479 HKHGMQLRRVG----PN 491
            ::|: :|||||
Db  486 -RNGIMVRRVGGDNDPN 501
```

RESULT 3
T06585
ammonium transporter 2 - tomato
C;Species: Lycopersicon esculentum (tomato)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 20-Jun-2000
C;Accession: T06585
R;Lauter, F.R.; Ninnemann, O.; Frommer, W.B.
submitted to the EMBL Data Library, January 1996
A;Description: Cloning and characterization of the ammonium transporter LeAMT2 from toma
A;Reference number: Z15776
A;Accession: T06585
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-514 <LAU>
A;Cross-references: EMBL:X95098; PIDN:CAA64475.1
A;Experimental source: seedling; root hair
C;Genetics:
A;Gene: AMT2
C;Superfamily: ammonium transport protein amt1

```
Query Match      74.0%;  Score 1953;  DB 2;  Length 514;
Best Local Similarity  75.3%;  Pred. No. 7.7e-141;
Matches 379;  Conservative 43;  Mismatches 71;  Indels 10;  Gaps 7;

Qy    2 SLPACPAEQLAQLLGPNTTDASAAASLICGHFAAVDSKFVDTAFAVDNTYLLFSAYLVFS 61
            | :||  |::| | |  ::|| |: || : | |:|| : :|||| : :||||||||||||:
Db    3 SAMTCSAAELFPHLG-SSANATARAAEFICSRFSAVSEYLTNTTYAVDTTYLLFSAYLVFA 61

Qy   62 MQLGFAMLCAGSVRAKNTMNIMLTNVLDAAAGGLFYYLFGFAFAFGSPSNGFIGKHFFGL 121
            ||||||||||||||||||||||||||||||||| |||||||||||| ||||||||||||||
Db   62 MQLGFAMLCAGSVRAKNTMNIMLTNVLDAAAGGFSYYLFGFAFAFGAPSNGFIGKHFFGL 121

Qy  122 KDIPSSSYDYSYFLYQWAFAIAAAGITSGSIAERTQFVAYLIYSSFLTGFVYPVVSHWFW 181
            |: || :::|||||||||||||||||||||||||||||||||||||||||||||:|||||
Db  122 KEFPSQAFDYSYFLYQWAFAIAAAGITSGSIAERTQFVAYLIYSSFLTGFVYPIVSHWFW 181

Qy  182 SPDGWASAFKITD--RLFSTGVIDFAGSGVVHMVGGIAGLWGALIEGPRMGRFDHAGRAV 239
            | ||||| :  ||   :| :|||||||||||||||||||||||:|||||:|||| : || |
Db  182 SGDGWASASK-TDGNLLLRFGVIDFAGSGVVHMVGGIAGLWGAFIEGPRIGRFDRSGRSV 240

Qy  240 ALRGHSASLIVVLGTFLLWFGWYGFNPGSFNKILLTYGNS--GNYYGQWSAVGRTAVTTTL 297
            ||||||||| ||||||||||||||||||||  |:| || :    :::||||:||||||||||
Db  241 ALRGHSASLVVLGTFLLWFGWYGFNPGSFLTILKSYDHTIRGTYYGQWSAIGRTAVTTTL 300

Qy  298 AGSTAALTTLFGKRVISGHWNVTDVCNGLLGGFAAITAGCSVVEPWAAIVCGFVASIVLI 357
            || |||||||| ||:: ||||| |||||||||||||::||: ||||||||||||:|| |||
Db  301 AGCTAALTTLFCKRLLVAHWNVDVCNGLLGGFAAITSGCAAVEPWAAIVCGFIAAWLI 360

Qy  358 ACNKLAEKVKFDDPLEAAQLHGCGCGTWGVIFTALFAKKEYVKEVY-GL-GRAHGLLMGGG 415
               ||| ||||||||||||| | | :|||:||:|||||| :| :| ::: |||:||:|||
Db  361 GFNALAAKLKYDDPLEAAQLHGGCGGSWGIIFTGLFAKKEVNEVYPGFPNRPYGLFMGGG 420

Qy  416 GKLLAAHVIQILVIAGWVSATMGPLFWGLNKLKLLRISSEDELAGMDMTRHGGFAYAY-- 473
            |||| |:|||:|:| |||||:||||| : |::|||||| ||: ||||::|||||||| | 
Db  421 GKLLGAQVIQVVIIGWVSVTMGPLFYLLHKFKLLRISRDDETAGMDLTRHGGFAYIYHD 480

Qy  474 EDDETHKGMQLRRVGPNASSTP 496
            ||  ::  : : ||| | |:|||
Db  481 EDBGSSMPGFKMIRVEPTNTSTP 503
```

RESULT 4
A96671
Ammonium transporter ATM1,2 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 02-Mar-2001
C;Accession: A96671
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, X.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: A96671
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-514 <STO>
C;Genetics:
A;Gene: F13O11.9
A;Map position: 1
C;Superfamily: ammonium transport protein amt1

Query Match   72.7%;   Score 1919;   DB 2;   Length 514;
Best Local Similarity   73.7%;   Pred. No. 3e-138;
Matches 370;   Conservative 48;   Mismatches 74;   Indels 10;   Gaps 6;

```
Qy    6 CPAEQLAQLL--GPNTTDASAAAASLICGHFAAVDSKFVDTAFAVDNTYLLFSAYLVFSMQ  63
          |||||  |     ||||:| ||:   ::|||||||||||||||| ||||||||||||||||
Db    8 CSAVDLSALLSSSNSTSSLAAATFLCSQISNISNKLSDTTYAVDNTYLLFSAYLVFAMQ  67

Qy   64 LGFAMLCAGSVRAKNTMNIMLTNVLDAAAGGLFYLFGFAFGSPSNGFIGKH--FFGL  121
           |||||||||||||||||||||||||||||||||  :|||| ||| :|||| |  |:||
Db   68 LGFAMLCAGSVRAKNTMNIMLTNVLDAAAGAISYYLFGFAFGTPSNGFIGRHSFFAL  127

Qy  122 KDIPS-SSYDYSYFLYQWAFAIAAAGITSGSIAERTQFVAYLIYSSFLTGFVYPVVSHWF  180
          ||| | |:::|||||||||||||||||||||||||||||||:||||:|||||||:|||:|
Db  128 SSYPERPGSDFSFFLYQWAFAIAAAGITSGSIAERTQFVAYLIYSTFLTGFVYPTVSIWF  187

Qy  181 WSPDGWASAFKITDR-LFSTGVIDFAGSGVVHMVGGIAGLWGALIEGPRMGRFDHAGRAV  239
          |||||||||| || | ||| ||||||||||||||||||||||||||||| | |  |||| 
Db  188 WSSDGWASASRSDNNLLFGSGAIDFAGSGVVHMVGGIAGLCGALVEGPRIGRFDRSGRSV  247

Qy  240 ALRGHSASLVVLGTFLHFGWYGFNPGSFNKILLTYGNSGNYYGQWSAVGRTAVTTLLAG  299
          ||||||||||||||||  |||||||||||  |||||||||||  ||||||||||||| | 
Db  248 ALRGHSASLVVLGTFLLWFGWYGFNPGSFLTILKGYDKSRPYYGQWSAVGRTAVTTLSG  307

Qy  300 STAALTTLFGKRVISGHWNVTDVCNGLLGGFAAITAGCSVVEPWAAIVCGFVASIVIIAC  359
          ||||||||||:|| ::  ||||||||||||||||| |  :|:||||||||||||| |  |
Db  308 CTAALTTLFSKRLLAGHWNVIDVCNGLLGGFAAITSGCAAVEPWAAIVCGFVASWVLIGF  367

Qy  360 NKLAEKVKFDDPLEAAQLHGGCGTWGVIFTALFAKKEYVKEVVGLGRAHGLLMGGGGKLL  419
          |||| |||||||||||||||||| ||||||| | :|||  |  |  | || ||||| |||
Db  368 NLLAKKLKYDDPLEAAQLHGGCGAWGLIFTGLFARKEYVNEIYSGDRPYGLFMGGGGKLL  427

Qy  420 AAHVIQILVIAGWVSATMGPLFWGLNKLKLRISSEDELAGMDMTRHGGFAYAYEDDE--  477
          |  |:||:| ||||:|||||||| |: |:|||||||:|||||| ||||||||| :
Db  428 AAQIVQIIVIVGWVTVTMGPLFYGLHKMNLRISAEDEMAGMDMTRHGGFAYAYNDEDDV  487

Qy  478 -THKGHGMQLRRVGPNA-SSTPT  497
           ||| | | |:| ||| ||||||
Db  488 STKPWGHFAGRVEPTSRSSTPT  509
```

RESULT 5
T03441
probable ammonium transport protein 1 - rice

C;Species: Oryza sativa (rice)
C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 24-Nov-1999
C;Accession: T03441
R;von Wiren, N.; Bergfeld, A.; Ninnemann, O.; Frommer, W.B.
submitted to the EMBL Data Library, April 1997
A;Description: OsAMT1-1, a putative ammonium transporter cDNA from rice.
A;Reference number: Z14948
A;Accession: T03441
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-533 <VON>
A;Cross-references: EMBL:AF001505; NID:g2160781; PID:g2160782
A;Experimental source: cv. Nipponbare, root
C;Genetics:
A;Gene: AMT1
C;Superfamily: ammonium transport protein amt1

Query Match   67.9%;   Score 1792.5;   DB 2;   Length 533;
Best Local Similarity   71.6%;   Pred. No. 1.3e-128;
Matches 361;   Conservative 39;   Mismatches 77;   Indels 27;   Gaps 11;

```
Qy    3 LPACPAEQLAQLLGPNTTDASAAASLICGHFAAVDSKFVDTAFAVDNTYLLFSAYLVFSM  62
          :  |  |  |   |||||||| ||:||||||||||||  ||:| |||||||||||||||
Db    1 MATCAAD-LAPLLGP---VAANATDYLCNRFA-------DTTSAVDATYLLFSAYLVFAM  49

Qy   63 QLGFAMLCAGSVRAKNTMNIMLTNVLDAAAAGLFYLFGFAF---AFGSPSNGFIGKHFF  119
          |||||||||||||||||||||||||||||| || |||||||    |||| :|||||| |
Db   50 QLGFAMLCAGSVRAKNTMNIMLTNVLDAAAGALFYLFGFASLRDCLRTPSNGFIGKQFF  109

Qy  120 GLKDIPS-SSYDYSYFLYQWAFAIAAAGITSGSIAERTQFVAYLIYSSFLTGFVYPVVSH  178
          ||| ::  |:|:|||| |||||||||||||||||||||||||||||||:|| |||||||
Db  110 GLKHMPARPGSDYDFFLPQWAFAIAAAGITSGSIAERTQFVAYLIYSAFLTGFVYPVVSH  169

Qy  179 WFWSPDGWASAFKIT-DRLFSTGVIDFAGSGVVHMVGGIAGLWGALIEGPRMGRFDHAGR  237
          | |:::||||| ||  ||||| |||||||||||||||||||||||||||| |||||||||
Db  170 WIWSADGWASASRTSGPLLFGSGVIDFAGSGVVHMVGGIAGLWGALIEGPRIGRPDHAGR  229

Qy  238 AVALRGHSASLVVLGTFLLWFGWYGFNPGSFNKILLTYGNSGNYYGQWSAVGRTAVTTL  297
          ||:|:||||||||||||||||||||||||||  |||| ||| ||||||||||||||||||
Db  230 SVALKGHSASLVVLGTFLLWFGWYGFNPGSFTTILKTYGPAGGINGQWSGVGRTAVTTTL  289

Qy  298 AGSTAALTTLFGKRVISGHWNVTDVCNGLLGGFAAITAGCSVVEPWAAIVCGFVASIVLI  357
          |||:||||||||||::||||||||||||||||||||||:  ::| |||||||||||| ||
Db  290 AGSVAALTTLFGKRLQTGHWNVVDVCNGLLGGFAAITAGGGVVDPWAAIICGFVSAWVLI  349

Qy  358 ACNKLAEKVKFDDPLEAAQLHGGCGTWGVIFTALFAKKEYVKEVVGLGRAHGLLMGGGGK  417
          |  :|:  ||||||||||||||||| | ||||||||| : |   || | ||| |  |||
Db  350 GLN--LAARLKFDDPLEAAQLHGGCGAWGILFTALFARQKYV-ERSTAPAAYGLFM-GRRK  406

Qy  418 LLAAHVIQILVIAGWVSATMGPLFWGLNKLKLRISSEDELAGMDMTRHGGFAYAYEDDE  477
          ||| |||||:|| |||: |||||||: |:|:||||||:|:||||||||||||||| |
Db  407 LLVGAVIQILVIFGWVSCTMGPLFYGLKKLGLLRISGRDETSGMDLTRHGGFAYVHDED  466

Qy  478 --THKH----GMQLR---RVGPNAS  493
            ||||    |:: |   |||| :
Db  467 EHDKSGVGGFMLRSAQTRVEPAAA  490
```

RESULT 6
H72379
ammonium transporter - Thermotoga maritima (strain MSB8)
C;Species: Thermotoga maritima
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C;Accession: H72379
R;Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey,
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;
C.M.
Nature 399, 323-329, 1999
A;Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
A;Reference number: A72200; MUID:99287316; PMID:10360571
A;Accession: H72379
A;Status: preliminary
A;Molecule type: DNA

A;Residues: 1-435 <ARN>
A;Cross-references: GB:AE001720; GB:AE000512; NID:g49980906; PIDN:AAD35487.1; PID:g49980906
A;Experimental source: strain MSB8
C;Genetics: TM0402
A;Gene: TM0402
C;Superfamily: ammonium transport protein amt1

```
Query Match          29.2%;  Score 771.5;  DB 2;  Length 435;
Best Local Similarity 38.2%;  Pred. No. 4.1e-51;
Matches 166;  Conservative 75;  Mismatches 157;  Indels 37;  Gaps 9;

Qy    42 DTAFAVDNTYLLFSAYLVFSMQLGFAMLCAGSVRAKNTMNIMLTNVLDAAAGGLFYYLFG 101
           ::  ||   :|:| |||  |||||||||||||||||||||||||||||||||| |:||
Db    26 DVGWSLDMWILISAALVFFMQAGFAMVESGFTRAKNTVNVLMKNLMDFAIGSVVFFIFG 85

Qy   102 FAFAFGSPSNGFIGKHFFGLKDIPSSS---YDYSYFLYQWAFAIAAAGITSGSIAERTQF 158
           :  :|||| ||| |||| | |  ::   | |: :||||||:|:|:||||||| |||:|
Db    86 YWIMF-------GKH--PLTFDPSTKEGLWDFAWWMFQMAFAGTAATIVSGAMAERTKF 135

Qy   159 VAYLIYSSFLTGFVYPVVSHWFWSPDGWASAFKITDRLFSTGVIDFAGSGVVHMVCGGIAG 218
           || |  :||||| |  |::|  | :  |||||||:|:|   | | | |||::|| ::||
Db   136 PAYLAYTGFITGIIYSVVGRWIWG-------GGW------LAQKGFIDFAGSTVVHSVGGWAA 185

Qy   219 LWGALIEGPRMGRFPDHAGRAVALRGHSASLVVLGTFLLWFGWYGFNPGSFNKILLTYGNS 278
           |||| ||||:||| ||||||||||||||||||||||||||||||||||| :|:|:|||||
Db   186 MIGASLLGPRFGKYDSQGNPKPIPGHNIPLAALGTPILWFGWFGFNGGS-----TLAGTN 240

Qy   279 GNYYGQWSAVGRTAVTTLAGSTAALTTLFGKRVISGHWNVTDVCNGLLGGFAAITAGCS 338
           | |   || :| || |:||| || :|| ||||||||||| || | |||||||:||||:|
Db   241 G------AIGMIILNTNLAAATGALAAMVTVWAKYGKPDASMTNNGALAGLVAITAPCA 293

Qy   339 VVEPWAAIVCGFVASIVLIACNKLAEKV-KFDDPLEAAQLHGGCGTWGVIFTALFAKKEY 397
           |:|| ||||||||:  | |  | | :|   |::|  |  | |||| |::|::|:||:|:
Db   294 VVSPVSSLIIGAICGVIVFAVEPFDKVLKIDDPVGAISVHGVNGAWGTLAVGLPAESKY 353

Qy   398 VKEVYGLGRAHGLLMGGGGKLLAAHVIQILVIAGWVSATMGPLFWGLNKLKLLRISSEDE 457
           |||||| |||:||:||||||| |||:|  |||||||| |||||||||||:|||||||:||
Db   354 AL-ASGMGDVNGLFFGGGGVHQLGVQFLGVVSVFAWTVVTSFIVFWFIKKTIGLRVDRDIE 412

Qy   458 LAGMDMTRHGGFAYA 472
           || :|||||||| :|
Db   413 LKGLDIEBHGMEGYA 427
```

RESULT 7
AC1930
ammonium transporter alr0990 [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
A;Accession: AC1930
C;Genetics:
A;Gene: alr0990
C;Superfamily: ammonium transport protein amt1
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
; Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001.
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A;Reference_number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AC1930
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-498 <KUR>
A;Cross-references: GB:BA000019; GB:BA000019; PIDN:BAB72947.1; PID:g17130336; GSPDB:GN00179
C;Genetics:
A;Gene: alr0990
C;Superfamily: ammonium transport protein amt1

```
Query Match          27.1%;  Score 715.5;  DB 2;  Length 498;
Best Local Similarity 34.4%;  Pred. No. 8.7e-47;
Matches 169;  Conservative 71;  Mismatches 184;  Indels 67;  Gaps 10;

Qy    24 AAASLICGHFAAVDSKFVDT----------AFAVDNTYLLFSAYLVFSMQLGFAMLC 70
           :::  || | || |     |              | || |||||||| ||| ||||:
Db    25 AIGSMVFAVFAPTIVQAVDTPTLESLSETTIKLQISIDTTWVLLSGFLVFFMQTGFAMLE 84
```

RESULT 8
A69468
ammonium transporter (amt-2) homolog - Archaeoglobus fulgidus
C;Species: Archaeoglobus fulgidus
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 29-Sep-1999
C;Accession: A69468
R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
; Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A;Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.
Smith, H.O.; Woese, C.R.; Venter, J.C.
A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo
A;Reference_number: A69250; MUID:98049343; PMID:9388475
A;Accession: A69468
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-468 <KLE>
A;Cross-references: GB:AE000982; GB:AE000782; NID:g2689305; PIDN:AAB89503.1; PID:g2689305
C;Superfamily: ammonium transport protein amt1

```
Query Match          27.0%;  Score 711.5;  DB 2;  Length 468;
Best Local Similarity 37.3%;  Pred. No. 1.6e-46;
Matches 160;  Conservative 62;  Mismatches 186;  Indels 21;  Gaps 6;

Qy    47 VDNTYLLFSAYLVFSMQLGFAMLCAGSVRAKNTMNIMLTNVLDAAAGGLFYLFGFAFAF 106
           |||||| ::|||||||||||| |||| |||||||:|:|||:| |||||||||||:||||
Db    37 VDFVWALICGFLVFMQAGFAMLEAGFSRAKNVANVMKNLMDFAVGSLAFFAVGFALMM 96

Qy   107 GSPSNGFIGKHFFGLKDIPSSSYDYS---YFLYQWAFAIAAAGITSGSIAERTQFVAYLI 163
           ||||||||||:| || |||||  |:|    ||| | || :|||||| |||||:|:| ||:
Db    97 GADWQGIAGTGWFL---AGESYDVSTELWFPMLVFRAATAATIVSGSIAERPKFSVYLV 153

Qy   164 YSSFLTGFVYPVVSHWFWSPDGWASAFKITDRL-FSTGVIDFAGSGVVHMVGGIAGLWGA 222
           || :|:| :| ||  |  ||||   |  :| |:|   | ||||| ||||| :  |:|
Db   154 YSAVVSAVIYPIYGHWLWG-GGWLSSSEFMVKLGGGYGALDFAGSGVVHALGGYIALAAV 212

Qy   223 LIEGPRMGRFDHAGRAVALRGHSASLVVLGTFLLWFGWYGFNPGSFNKILLTYGNSGNYY 282
           ::  ||||:|| ||||:|:|||||||||||||||||||||||||: ||:|||  ||||::
```

```
Db    213 MLLGPRLGKYDSDGNPRAIPGHNLAFAVIGTFILWFGWFGNAGS----------TLS 260
Qy    283 GQWSAVGRTAVTTLAGSTAALTLFGKRVISGHWNVTDVCNGLLGGFAAITAGCSVVEP 342
             |||  :|   :||: ||  | :  :|:|: | ::|
Db    261 AHELRVSIIASNTNLAAAGAVTAMAITWLRNGKPDVGMTCNGAVAGLVAITAPCAWVQP 320
Qy    343 WAAIVCGFVASIVLIACNKLAEKVKFDDPLEAAQLHGGCGTWGVIFTALFAKKEYVKEVY 402
             |::|:|:  ||:|    ||| |   :|  : |:|   | || ::|
Db    321 WSSVVIGTIAGFIATYGVWWLEKRGLDDVVGAIFVHGFSGTWGLLALGIFADGSYGLYAT 380
Qy    403 GLGRAHGLLMGGGKLLAAHVIQILVIAGWVSATMGPLFWGLNKLKLLRISSEDELAGMD 462
             ||||
Db    381 ESPLVTGLLYGNWG-FFIVQLISAIVNFAWAFGTGFALFWILKKVIGIRVSPEEEMLGLD 439
                      : |  ||  : | :   :| | ||| |   | :|::|   ||:|||:
Qy    463 MTRHGGFAY 471
             |||
Db    440 IAEHAAVAY 448


RESULT 9
T15414
hypothetical protein C05E11.4 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 18-Feb-2000
C;Accession: T15414
R;Geisel, C.
submitted to the EMBL Data Library, April 1996
A;Description: The sequence of C. elegans cosmid C05E11.
A;Reference number: 218347
A;Accession: T15414
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-534 <GEI>
A;cross-references: EMBL:U53338; NID:g1255840; PID:g1255843; PIDN:AAA96191.1; GSPDB:GN00
C;Experimental source: strain Bristol N2; clone C05E11
C;Genetics:
A;Gene: CESP:C05E11.4
A;Map position: X
A;Introns: 43/1; 151/1; 182/3; 229/3; 339/2; 394/3; 423/3; 490/1
C;Superfamily: ammonium transport protein amt1

Query Match          26.8%;  Score 706;  DB 2;  Length 534;
Best Local Similarity 36.0%;  Pred. No. 4.9e-46;
Matches 164;  Conservative 74;  Mismatches 185;  Indels 32;  Gaps 9;

Qy     41 VDTAF--AVDNTYLLFSAYLVFSMQLGFAMLCAGSVRAKNTMNIMLTNVLDAAAGGLFYY 98
Db     21 LETGFYENVNSFFLCSMALIIIFMQCGFAYLEAGAVRSKNTTNILIHKNLLDSCICIIGYW 80
Qy     99 LFGFAFAFGSPSNG---FIGKHFFGLKDIPSSSYDYSFYLYQWAFAIAAAGITSGSIABR 155
Db     81 AIGWALAYGDSGEGVNLFVGHSQFFL----SGFSDYPRFFFQVVFSATAATIVSGAVAER 136
Qy    156 TQFVAYLIYSSFLTGFVYPVPVVSHWFWSPDGWASAFKITDRLFSTGVIDFAGSGVVHMVGG 215
Db    137 CEFITVYTYCTVISTFIYPVLTHWGWTENGW-MAKGIISGIIDIKYDDFAGSGLVHLCGG 195
Qy    216 IAGLWGALIEGPRMGRF--DHAGRAVALRGHSASLVVLGTFLLWFGWYGFNPGSFNKILL 273
Db    196 SISFLRAWIMGPRIGKFPDDEDDESDEIIGHSVPFTALGGFILMFGFLAFNGGSVASI-- 253
Qy    274 TYGNSGNYYGQWSAVGRTAVTTLAGSTAALTLFGKRVISGHWNVTDVCNGLLGGFAAI 333
Db    254 ------SHAGDGHTVALAMINTILSGFAALIIYLGVHYYQHGKWTLLLTINACLSGMVAA 307
Qy    334 TAGCSVVEPWAAIVCGFVASIVLIACNKLAEKVKFDDPLEAAQLHGGCGTWGVIFTALFA 393
Db    308 CAGCNKMEPWACIWVGLGAGLIYLAFSKLMIRLKIDDPLDAFAVHAGGFWGLMSSSIIS 367
Qy    394 KKEYVKEVYGLGRAHGLLMGGGKLLAA------HVIQILVIAGWVSATMGPLFWGLNKL 447
Db    368 ---HGGVAYALADAVSGAKNSGDHLTQAFAQLGWQMICALAIIAWSLGVMLPIFWILKKT 424
```

```
Qy    448 KLLRISSEDELAGMDMTRHGGFAYAYEDDETHKHG 482
Db    425 GKLRVSEEVEINGLDVFKHGEMAYPL--RAYGHG 456


RESULT 10
S76687
hypothetical protein - Synechocystis sp. (strain PCC 6803)
C;Species: Synechocystis sp.
C;Variety: PCC 6803
C;Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
C;Accession: S76687
R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
o, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasudi
DNA Res. 3, 109-136, 1996
A;Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
s.
A;Reference number: S74322; MUID:97061201; PMID:8905231
A;Accession: S76687
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-507 <KAN>
A;Cross-references: EMBL:D64004; GB:AB001339; NID:g1001701; PIDN:BAA10631.1; PID:g12084(
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C;Superfamily: ammonium transport protein amt1

Query Match          26.6%;  Score 703;  DB 2;  Length 507;
Best Local Similarity 36.2%;  Pred. No. 7.9e-46;
Matches 160;  Conservative 63;  Mismatches 173;  Indels 46;  Gaps 8;

Qy     51 YLLFSAYLVFSMQLGFAMLCAGSVRAKNTMNIMLTNVLDAAAGGLFYYLFGFAPAFGSPS 110
Db     92 WILIAAILVIFMNAGFGMLETGLCRQKNAVNILTKNLIVFALATIAYWAIGFSLMFGSSG 151
Qy    111 NGFI--------GKHF-FGLKDIPSSSYDYSFLYQWAFAIAAAGITSGSIAERTQFVAY 161
Db    152 NPFVGFGGFFLSGDHTNYGLSPFFPEGLPVAVFFLFQVAFSATAATIVSGAVAERIKFNEF 211
Qy    162 LIYSSFLTGFVYPVPVVSHWFWSPDGWASAFKITDRLFSTGVIDFAGSGVVHMVGGIAGLWG 221
Db    212 LIFSVLLVGIAYPITGHWVWDAGGW---------LYTMGFMDFAGSTVHSVGGWAALAG 262
Qy    222 ALIEGPRMGRFDHAGRAVALRGHSASLVVLGTFLLWFGWYGFNPGSFNKILLTYGNSGNY 281
Db    263 AFLLGPRLGKFVD-GRPGAIPGHNMGFAMLGCLIIWIGWFGFNPGS----------Q 308
Qy    282 YGQWSAVGRTAVTTLAGSTAALTLFGKRVISGHWNVTDVCNGLLGGFAAITAGCSVVE 341
Db    309 LAADQACAYIAVTTNLAASAGGLTATFTSWLKDGKPDLTMVINGVLAGLVGITAGCAGVS 368
Qy    342 PWAAIVCGFVASIVLIACNKLAEKVKFDDPLEAAQLHGGCGTWGVIFTALFAKKEYVKEV 401
Db    369 YWGSVIIGGIAGILVVYSVAFFDKIKIDDDPVGAISVHLVNGWGTLAVGFFNMEK----- 423
Qy    402 YGLGRAHGLLMGGGKLLAAHVIQILVIAGWVSATMGPLFWGLNKLKL-LRISSEDELAG 460
Db    424 ------GLFYGGGINQLIIQIVGLLAI-GAFTAIFSFVVWAILKQTMGIRVSGEEREMIG 475
Qy    461 MDMTRHGGPAYAYEDDETHKHG 482
Db    476 LDIGEHGMEAYTGFVKETDSFG 497


RESULT 11
S74801
ammonium transport protein - Synechocystis sp. (strain PCC 6803)
N;Alternate names: protein sll1017
C;Species: Synechocystis sp.
C;Variety: PCC 6803
C;Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
C;Accession: S74801
R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
o, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasudi
```

DNA Res. 3, 109-136, 1996
A;Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
s.
A;Reference number: S74322; MUID:97061201; PMID:8905231
A;Accession: S74801
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-442 <KAN>
A;Cross-references: EMBL:D90901; GB:AB001339; NID:g1651897; PIDN:BAA16952.1; PID:g165202
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C;Genetics:
A;Gene: amt1
C;Superfamily: ammonium transport protein amt1

Query Match   25.7%; Score 678.5; DB 2; Length 442;
Best Local Similarity 35.9%; Pred. No. 4.9e-44;
Matches 168; Conservative 61; Mismatches 152; Indels 87; Gaps 15;

Qy   38 SKFVDTAPAVDNTYLLFSAYLVFSMQLGFAMLCAGSVRAKQTMNIMLTNVLDAAAGGLFY 97
Db   32 TSIAEVTYAINNLFLLAAAVLVLFMQAGFAMLEAGLSSHKNTVNVLFKNTFDVCVGVLLY 91

Qy   98 YLFGFAFAFGSPS--NGFIGKHFG----LKDIPSSSYDYSYFLYQWAFAIAAAGITSGS 151
Db   92 FLFGYSLMYGENPVLGGFGWGGFGITNNLDNVEGLSPQVDW-LFQAAFAATAATTVSGA 150

Qy 152 IAERTQFVAYLIYSSFLTGFVYPVSHHWFWSPDGWASAFKITDRLFSTGVIDFAGSGVH 211
Db 151 VMGRMYFKAYLIYSAVITGLVYPISGHWKWG------GGW------LDKL--GFHDFAGSLLVH 200

Qy 212 MVGGIAGLWGALIEGPRMGRFDHAGRAVALRGH----SASLVVLGTFLLMFGWYGFNPGS 267
Db 201 SVGGFAALAAVVVMGPRIGRFE--GNKINSLGYGQITSSS---LGVFILWVGWYGFNPGS 255

Qy 268 FNKILLTYGNSGNYYGQWSAVGR-------TAVTTTLAGSTAALTTLFGKRVISGHW-- 317
Db 256 ---------------------QLAFVGALNTNTTMLIAVNTTLSSAAAGGLAAL-----AFDWIT 293

Qy 318 ----NVTDVCNGLLGGFAAITAGCSVVEPWAAIVCGFVASIVLIACNKLAEKVKFDDP 371
Db 294 ENKRKPNLLVTLNGILGLVGITPAGCDTVSNWSAIAIGVVAGILSVLGTKLLDRLRIDDG 353

Qy 372 LEAAQLHGGCGTWGVIFTALFAKKEYVKEVYGLGRAHGLLMGGGGKLLAAHVIQILVIAG 431
Db 354 ---LSAQIVGSIVIPF 394

Qy 432 WVSATMGPLFWGLNKLKLLRISSEDELAGMDMTRHG----GFAYAYED 475
Db 395 WAFITMFLFYVMDLWGILRVKPSQEKVGLDIVEHGQTEKGVEIAFED 442

RESULT 12
AD1930
ammonium transporter alr0991 [imported] - Nostoc sp. PCC 7120
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C;Accession: AD1930
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AD1930
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-518 <KUR>
A;Cross-references: GB:BA000019; PIDN:BAB72948.1; PID:g17130337; GSPDB:GN00179
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: alr0991
C;Superfamily: ammonium transport protein amt1

Query Match   25.2%; Score 665; DB 2; Length 518;
Best Local Similarity 34.3%; Pred. No. 6.2e-43;
Matches 154; Conservative 77; Mismatches 170; Indels 48; Gaps 11;

Qy   46 AVDNTYLLFSAYLVFSMQLGFAMLCAGSVRAKNTMNIMLTNVLDAAAGGLFYLFGFAFA 105
Db   76 ALDTLWVAIAFLVFFMNAGFGMLETGFCRQKNAVNVLAKNIIVFALATVAFWAIGFGLM 135

Qy 106 FGSPSN--GFIGKHFFGLKDIPSSSYDY---------SYFLYQWAFAIAAAGIT 148
Db 136 FGDGNDFIGFGNGLFLSGVDNSPATGDAYKGVFSALSWAGVPLAAKFLFQLVFAGTAATIV 195

Qy 149 SGSIAERTQFVAYLIYSSFLTGFVYPVSHHWFWSPDGWASAFKITDRLFSTGVIDFAGSG 208
Db 196 SGAVAERIKFVDFLIFSLLLVGIAYPITGHWIWGA-GW----------LAKAGFWDFAGST 245

Qy 209 VVHMVGGIAGLWGALIEGPRMGRFDHAGRAVALRGHSASLVVLGTFLLWFGWYGFNPGSF 268
Db 246 VVHSVGGWAALMGAAFLGPRIGKYQDK-QIVALPGHNMSIATLGCLILWLGWFGFNPGS- 303

Qy 269 NKILLTYGNSGNYYGQWSAVGRTAVTTTLAGSTAALTTLFGKRVISGHWNVTDVCNGLLG 328
Db 304 --VMAADPN--------AITHIALTTNMAGAVGGIAATAWLYLGKPDLSMIINGILA 352

Qy 329 GFAAITAGCSVVEPWAAIVCGFVASIVLIACNKLAEKVKFDDPLEAAQLHGGGTWGVIF 388
Db 353 GLVGITASCAYVSIPSSIIIGLIAGVIVPSVTFPDKLGIDDPVGATSVHLVCGVWGTIA 412

Qy 389 TALFAKKEYVKEVY--GLGRAHGLLMGGGGKLLAAHVIQILVIA--GWVSATMGPLFWGL 444
Db 413 VGLWSVGPGVYSWYGEGLGPTKGLFAGGG---LGQLITQFLGAAAVGGMTVLVSSIFWVV 469

Qy 445 NKLKL-LRISSEDELAGMDMTRHGGFAYA 472
Db 470 LKATLGIRVTREEBLEGLDIGEHGMEAYS 498

RESULT 13
B84129
ammonium transporter BH3834 [imported] - Bacillus halodurans (strain C-125)
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C;Accession: B84129
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: B84129
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-421 <STO>
A;Cross-references: GB:AP001520; GB:BA000004; NID:g10176401; PIDN:BAB07553.1; GSPDB:GN00
A;Experimental source: strain C-125
C;Genetics:
A;Gene: BH3834
C;Superfamily: ammonium transporter nrgA

Query Match   24.8%; Score 655; DB 2; Length 421;
Best Local Similarity 32.9%; Pred. No. 2.9e-42;
Matches 153; Conservative 69; Mismatches 175; Indels 68; Gaps 13;

Qy   41 VDTAFAVDNTYLLFSAYLVFSMQLGFAMLCAGSVRAKNTMNIMLTNVLDAAAGGLFYLF 100
Db   1 MDSIALMDNVWIMLCAVLVLLMQGGFILLEAGSTRMKNAGHIAGKTIFTVGIASLVTWAV 60

Qy 101 GFAPAFGSPSNGFIG--KHFFGLKDIPSSSYDYSY----------FLYQWAFAIAAAGIT 148
Db   61 GWGFTYGB-GNAFIGLSDFFFG----------DYSTAEEGLVGSVDFFFQLMFAAILTIA 110

Qy 149 SGSIAERTQFVAYLIYSSFLTGFVYPVSHHWFWSPDGWASAFKITDRLFSTGVIDFAGSG 208
Db 111 FGGFAERAKLSVYILFAVLFSAFVYPIVAHWIWG-DGW----------LADLGKQDFAGST 160

Qy 411 LMGGGGKLLAAHVIQILVIAGWV----------SATMGPL-FWGLNKLKL-LRISSEDE 457
              |||:   ::|| :|| ||||::| :||:
Db 390 LYSWSFELL----------WVQLQCTAAILIYSATTGFLALFLISKSPIGLRVTDYEE 437

Qy 458 LAGMDMTRHG 467
              :|::||
Db 438 QIGADVIEHG 447

RESULT 15
S76017
hypothetical protein - Synechocystis sp. (strain PCC 6803)
C;Species: Synechocystis sp.
A;Variety: PCC 6803
C;Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
C;Accession: S76017
R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
o, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A;Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
s.
A;Reference number: S74322; MUID:97061201; PMID:8905231
A;Accession: S76017
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-541 <KAN>
A;Cross-references: EMBL:D64006; GB:AB001339; NID:g1001291; PIDN:BAA10864.1; PID:g100137
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C;Superfamily: ammonium transport protein amt1

Query Match   21.6%; Score 569.5; DB 2; Length 541;
Best Local Similarity   33.2%; Pred. No. 1.2e-35;
Matches 153; Conservative 58; Mismatches 191; Indels 59; Gaps 12;

Qy   46 AVDNTYLLFSAYLVFSMQLGFAMLCAGSVRAKNTMNIMLTNVLDAAAGGLFYLFGFAFA 105
              ::|:||||:|||||||||: |||:||:::|:|:||||: |||||:||:::
Db   3 SIDTLWILLCAGLVFFWQAGFMCLESGLTRSKNSINVAIKNFADFGISVALFWSFGFSIM 62

Qy 106 FGSPSNGFIGKHFFGLKDIPSSSYDYSYFLYQWAFAIAAAGITSGSIAERTQFVAYLIYS 165
              |||:|||||| |||||:||| :|:|: |::||||| |||||||:|||||||||:|
Db   63 FGLSQGGWWGTG-YSFVDVGGEPTLAVFLFQAMFCGTATTIISGAAAERLKFSAYLLVA 121

Qy 166 SFLTGFVYPVVSHWFW---------------SPDGWASAAFKITDRLFSTGVIDFAGSGVVHMVGG 215
              ::|: :||||||: | :| : ||:::||::::|:|||||||::||
Db 122 GLASGLIYPLFGDWAWNGLATVAGIETTGGW----------LENLGFRDFAGSTVHSVGA 172

Qy 216 IAGLWGALIEGPRMGRFPDHAGRAVALRGHSASLVVLGTFLLWFGWYGFNPGSFNKILLTY 275
              || :| |:|||:|||:||||||||:|||||| ||||||| ||:|||:| :||||||:
Db 173 WIGLATIIVVGPRQGRFPKTGKTLKIQGSNMPFSVLGTLILWFGWLGFNGGS------TF 226

Qy 276 GNSGNYYGQWSAVGRTAVTTTLAG---STAALTTLFGKRVISGHWNVTDVCNGLLGGFA 331
              ||:|:|||:|| |:|::|||||| |:| :::|:::|||:|:::|||| |||
Db 227 GLTPEVPG-------IMVNTVLAGVGGMLMAGLISLLQDKMI----QVEPIMNGSLAGLV 275

Qy 332 AITAGCSVVEPAAAIVCGFVASIVLIACNKLAEKVKFPDPLEAAQLHGGCGTWGVIFTAL 391
              || ::|||:||||:|||||||:| :|| |||||| :::| ||:||:||:||||:||||
Db 276 AITASANVMTPIAMVIGATGSAIAYLVGKKMLHWGVDDAVDAVAVHGGAGVWGTLCVGL 335

Qy 392 FAKKEYVKEVYGLGRAHGLLMGGGGKLIAAHVIQILVIA-GWVSATMGPLFWGLNKLKLL 450
              |:|:::: :|||:|| ||||||||||||||||||||:|| : |:::|:||| :||:
Db 336 FGQLPLVDT--GLNRWQQC----GVQLLGIGVCTIWAFGLAWVFITL-----LNRVFAL 383

Qy 451 RISSEDELAGMDMTRHGGFAYAYE-----DDETHKGGMQLR 486
              ||:: :|:|||:||:|| : :| :||:::|:
Db 384 RISPEDEEIGLNVSEHQATTETYELFQVMDRQAKTHDLSLR 424

Search completed: March 9, 2004, 11:56:07
Job time : 22 secs

Qy 209 VVHMVGGIAGLWGALIEGPRMGRFDHAGRAVALRGHSASLVVLGTFLLWFGWYGFNPGSF 268
              |||| :: :| |:|||:||:||||||||:|||||||||:||||| ||:|||:|
Db 161 VVHLTGAMGALAATILLKPRIGKYNKDGSMDIQGRNQVYALGVLLWGVNFGFNAGST 220

Qy 269 NKILLTYGNSGNYYGQWSAVGRTAVTTLLAGSTAALTTLFGKRVISGHWNVTDVCNGLLG 328
              :||||||:|:|:|||:|| |:|::|||||| | :::||:::|||:|:::|||| |
Db 221 LEVADAF----------FGYVALNTQLAAAGAVAAMFIAWALSGKADVPTLNGALA 268

Qy 329 GFAAITAGCSVVEPAAAIVCGFVASILVIACNKLAEKVKFPDPLEAAQLHGGCGTWGVIF 388
              | ::|| ::|||:||||:|||||||: :|| |||||| :::| ||:||:||:||||:
Db 269 GLVAITASCAFVAPWAAVVIGVIGGLIVFSMKFFPDKAKIDDPIFALSVHGVAGVVGWTLS 328

Qy 389 TALFAKKEYVKEVYGLGRAHGLLMGGGGKLLAAHVIQILV-----IAGWV----SATM 437
              |||||:::: :|||:|| |||||||||||||||||||||| :| |:::
Db 329 TGFFATPE-LAEMNG-GQA-GLFYGGGFAQLGVQTLSVVACGVFAFIASYVLLLVCKSVL 385

Qy 438 GPLFWGLNKLKLLRISSEDELAGMDMTRHGGFAYAYEDDETHKHG 482
              |::||| :||: |:::||:|:|||:||:|| : :|||||:||
Db 386 GG--------LRVTEEEEIIGLDLSEHGSGYPBSMPGSEKSG 420

RESULT 14
T23804
hypothetical protein M195.3 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T23804
R;Burton, J.
submitted to the EMBL Data Library, October 1995
A;Reference number: Z19801
A;Accession: T23804
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-622 <WIL>
A;Cross-references: EMBL:Z66498; PIDN:CAA91293.1; GSPDB:GN00020; CESP:M195.3
A;Experimental source: clone M195
C;Genetics:
A;Gene: CESP:M195.3
A;Map position: 2
A;Introns: 18/1; 55/1; 152/1; 315/1; 379/3; 421/3; 465/2; 508/1; 554/1

Query Match   22.0%; Score 579.5; DB 2; Length 622;
Best Local Similarity   31.8%; Pred. No. 2.4e-36;
Matches 156; Conservative 86; Mismatches 169; Indels 79; Gaps 19;

Qy   14 LLGPNTT--DASAAASLICGHFAAVD-SKFVDTAFAVDNTYLLFSAYLVFSMQLGFAMLC 70
              ||||: : |: ::| : :| ::|::| |:||||||||:|||||||||:
Db   1 MAGPEGSIFNASAMQIVQIHHYAEGSVTPEVDKLYQDDAVWIISSFIIFTMHSGFGLLE 60

Qy   71 AGSVRAKNTMNIMLTNVLDAAAGGLFYLFGFAFG------SPSNGFIGKHF------ 118
              ||:|:||: |:|:||:|||||:||:||||||: |||||||
Db   61 SGSVSAKDEVNIMVKNVVDVVFGGLSYWSCGFSYGDIPEWRNPYVGF-GKFFYDPTRD 119

Qy 119 FGLKD-IPSSSYDYSYFLYQWAFAIAAAGITSGSIAERTQFVAYLIYSSFLTGFVYPVVS 177
              |||: ||||||:||:|||| |:||:::| :|||:::|||::||| ||:||:| :||||
Db 120 YGTRETINQEGWSYASFLFQLSLATTASTIVSGAVAERAKLKSYILLGCIVI--LIQALPA 178

Qy 178 HWFWSPDGWASAFKITDRLFSTGVIDFAGSGVVHMVGGIAGLWGALIEGPRMGRFDHAGR 237
              |::|||| :: ||:::||::::|:|||||||::||:::|| :| |:|||:||::||
Db 179 HWVWDKEG---------VFYKKGVVDFAGCSAVHLVGGIIGLLATVFLKPRRNRFNEDS- 228

Qy 238 AVALRGHSASLVVLGTFLLWFGWYGFNPGSFNKILLTYGNSGNYYGQWSAVGRTAVTTTL 297
              ||||:|||||||| |||||| :|||||:| :|||||||:|:|:|||:|| |:|:::||
Db 229 --VHQMSSPTNALLGTFLLWWGWFGINAGS------VWGITG---GRWRLGARAVAATM 277

Qy 298 A---GSTAALTTLFGKRVISGHWNVTDVCNGLLGGFRAITAGCSVVEPWAAIVCGFVASI 354
              :   :|: | :||:|:::|||:|:::|||| || || ::|||:||:||:||||||:
Db 278 ASIGGGATAITTISFVK---TKKLQVNFLINGILSSIVSITAICAVSRPWHALVIGSISSV 334

Qy 355 VLIACNKLAEKVKFPDPLEAAQLHGGCGTWGVIFTALFAKKEYVKEVYGLGRA---HGL 410
              ||:|| |||||| :::| ||:||:||:||||:|||||||:::: :|||:|||| |||
Db 335 FSIAVLPLLDRLHIDDPVGIVPIHLTSSIWGMIAVGIFCEEDKY-----LGSATNNRSGL 389

OM protein - protein search, using sw model

```
                Copyright    GenCore version 5.1.6
                             Copyright (c) 1993 - 2004 Compugen Ltd.

Run on:         March  9, 2004, 11:52:49 ; Search time 22 Seconds
                                           (without alignments)
                                           1173.318 Million cell updates/sec

Title:          US-10-033-109-4
Perfect score:  2639
Sequence:       1 MSLPACPAEQLAQLLGPNTT.........HGMQLRRVGPNASSTPTTDE 500

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters:      389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :    Issued_Patents_AA:*
              1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
              2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
              3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
              4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
              5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
              6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                                    SUMMARIES
```

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2129.5	80.7	501	4	US-08-635-967-2	Sequence 2, Appli
2	488	18.5	470	4	US-09-328-352-5397	Sequence 5397, Ap
3	468	17.7	433	4	US-09-543-681A-6105	Sequence 6105, Ap
4	450	17.1	439	4	US-09-489-039A-12960	Sequence 12960, A
5	442	16.7	455	4	US-09-252-991A-23745	Sequence 23745, A
6	441	16.7	431	4	US-09-328-352-4217	Sequence 4217, Ap
7	420.5	15.9	464	4	US-09-252-991A-18525	Sequence 18525, A
8	406.5	15.4	437	4	US-09-134-001C-3907	Sequence 3907, Ap
9	203.5	7.7	361	4	US-09-134-000C-5907	Sequence 5907, Ap
10	193	7.3	223	4	US-09-107-532A-3674	Sequence 3674, Ap
11	182	6.9	224	4	US-09-107-532A-4102	Sequence 4102, Ap
12	136	5.2	443	4	US-09-489-039A-9335	Sequence 9335, Ap
13	121.5	4.6	490	4	US-09-489-039A-9610	Sequence 9610, Ap
14	119.5	4.5	517	4	US-09-252-991A-19322	Sequence 19322, A
15	119	4.5	557	4	US-09-252-991A-18216	Sequence 18216, A
16	117.5	4.5	472	3	US-09-354-129-2	Sequence 2, Appli
17	117.5	4.5	472	4	US-09-504-357-2	Sequence 2, Appli
18	113	4.3	471	4	US-09-252-991A-22393	Sequence 22393, A
19	111.5	4.1	481	4	US-09-252-991A-27157	Sequence 27157, A
20	109.5	4.1	470	4	US-09-328-352-6912	Sequence 6912, Ap
21	108	4.1	1216	4	US-09-134-000C-5130	Sequence 5130, Ap
22	107.5	4.0	391	4	US-09-489-039A-9791	Sequence 9791, Ap
23	106.5	4.0	306	4	US-09-328-352-5398	Sequence 5398, Ap
24	106.5	4.0	472	4	US-09-489-039A-13479	Sequence 13479, A
25	106.5	4.0	678	4	US-09-252-991A-20202	Sequence 20202, A
26	104	3.9	463	4	US-09-252-991A-29995	Sequence 29995, A
27	104	3.9	656	4	US-09-543-681A-7550	Sequence 7550, Ap
28	104	3.9	657	4	US-09-252-991A-27682	Sequence 27682, A
29	103	3.9	452	4	US-09-543-681A-6544	Sequence 6544, Ap
30	102.5	3.9	339	4	US-09-107-532A-4904	Sequence 4904, Ap
31	102.5	3.9	351	4	US-09-489-039A-12237	Sequence 12237, A
32	102.5	3.9	417	1	US-08-553-888A-3	Sequence 3, Appli
33	102.5	3.9	457	4	US-09-489-039A-8037	Sequence 8037, Ap
34	102	3.9	467	4	US-09-489-039A-9693	Sequence 9693, Ap
35	102	3.9	512	4	US-09-540-236-3548	Sequence 3548, Ap
36	101.5	3.8	499	4	US-09-134-001C-5370	Sequence 5370, Ap
37	101	3.8	396	4	US-09-543-681A-5161	Sequence 5161, Ap
38	99.5	3.8	352	4	US-09-489-039A-8296	Sequence 8296, Ap
39	99	3.8	513	4	US-09-489-039A-12877	Sequence 12877, A
40	98.5	3.7	459	4	US-09-252-991A-22668	Sequence 22668, A
41	98.5	3.7	482	4	US-09-328-352-6084	Sequence 6084, Ap
42	98	3.7	389	4	US-09-489-039A-7563	Sequence 7563, Ap
43	98	3.7	455	4	US-09-489-039A-9942	Sequence 9942, Ap
44	97.5	3.7	439	4	US-09-172-952-14	Sequence 14, Appl
45	97	3.7	421	4	US-09-489-039A-7699	Sequence 7699, Ap

                                    ALIGNMENTS

```
RESULT 1
US-08-635-967-2
; Sequence 2, Application US/08635967
; Patent No. 6620610
; GENERAL INFORMATION:
;  APPLICANT: FROMMER, Wolf-Bernd
;  APPLICANT: NINNEMAN, Olaf
;  TITLE OF INVENTION:  DNA SEQUENCES FOR AMMONIUM TRANSPORTER,
;  TITLE OF INVENTION:  PLASMIDS, BACTERIA, YEASTS, PLANT CELLS AND PLANTS
;  TITLE OF INVENTION:  CONTAINING THE TRANSPORTER
;  NUMBER OF SEQUENCES: 2
;  CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Ostrolenk, Faber, Gerb & Soffen, LLP
;   STREET: 1180 Avenue of the Americas
;   CITY: New York
;   STATE: NY
;   COUNTRY: US
;   ZIP: 10036-8403
;  COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;  CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/635,967
;   FILING DATE: 29-APR-1996
;   CLASSIFICATION: 800
;  PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: PCT/EP 94/03499
;   FILING DATE: 24-OCT-1994
;  PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: DE P 43 37 597.9
;   FILING DATE: 28-OCT-1993
;  ATTORNEY/AGENT INFORMATION:
;   NAME: Meilman, Edward
;   REGISTRATION NUMBER: 24,735
;   REFERENCE/DOCKET NUMBER: P/951-120
;  TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (212) 382-0700
;   TELEFAX: (212) 382-0888
;   TELEX: 236925
;  INFORMATION FOR SEQ ID NO: 2:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 501 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
;  MOLECULE TYPE: protein
US-08-635-967-2

Query Match             80.7%;  Score 2129.5;  DB 4;  Length 501;
```

```
Best Local Similarity   80.9%;   Pred. No. 8.2e-220;
Matches  407;  Conservative  42;  Mismatches  39;  Indels  15;  Gaps  7;

Qy    5  ACPAEQLAQLLGPNTTDASAAASLICGHFAAVDSKFVDTAFAVDNTYLLFSAYLVFSMQL  64
            :|:||||||   |||||  :  |||||||  ||||||||||||||||||||||||
Db    2  SCSATDLAVLLGPN---ATAAANYICGQLGDVNNKFIDTAFAIDNTYLLFSAYLVFSMQL  58

Qy   65  GFAMLCAGSVRAKNTMNIMLTNVLDAAAGGLFYYLFGFAPAFGSPSNGFIGKHFFGLKDI  124
         |||||||||||||||||||||||||||||||||||||   |||||||||||| :||||||
Db   59  GFAMLCAGSVRAKNTMNIMLTNVLDAAAGGLFYYLFGYAFAFGSPSNGFIGKHYFGLKDI  118

Qy  125  PSSSYDYSYFLYQWAFAIAAAGITSGSIAERTQFVAYLIYSSFLTGFVYPVVSHWFWSPD  184
Db  119  PTASADYSNFLYQWAFAIAAAGITSGSIAERTQFVAYLIYSSFLTGFVYPVVSHWFWSVD  178

Qy  185  GWASAFKIT-DRLFSTGVIDFPAGSGVVHMVGGIAGLWGALIEGPRMGRFDHAGRAVALRG  243
Db  179  GWASPFRTDGDLLFSTGAIDFPAGSGVVHMVGGIAGLWGALIEGPRLGRFDNGGRAIALRG  238

Qy  244  HSASLVVLGTFLLWFGWYGFNPGSFNKILLTYGNSGNYYGQWSAVGRTAVTTLLAGSTAA  303
Db  239  HSASLVVLGTFLLWFGWYGFNPGSFNKILLVTY-ETGTYNGQWSAVGRTAVTTTLAGCTAA  297

Qy  304  LTTLFGKRVISGHWNVTDVCNGLLGGFAAITAGCSVVEPWAAIVCGFVASIVLIACNKLA  363
Db  298  LTTLFGKRILSGHHNVTDVCNGLLGGFAAITGGCSVVEPWAAIICGFVAAVLLGCNKLA  357

Qy  364  EKVKFDDPLEAAQLHGGCGTWGVIFTALFAKKEYVKEVYG--LGRAHGLLMGGGGKLLAA  421
Db  358  EKLKYDDPLEAAQLHGGCGAWGLIFTALFAQEKYLNQIYGNKPGRPHGLFMGGGGKLLGA  417

Qy  422  HVIQILVTAGWVSATMGPLFWGLNKLKLLRISSEDELAGMDTRHGGFAYAY-EDDETHK  480
Db  418  QLIQIIVITGWVSATMGTLFFILKKMKLLRISSEDEMAGMDTRHGGFAYMYFDDDESHK  477

Qy  481  HGMQLRRVGP------NASSTPT  497
Db  478  -AIQLRRVEPRSPSPSGANTTPT  499


RESULT 2
US-09-328-352-5397
; Sequence 5397, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 5397
; LENGTH: 470
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-5397

Query Match             18.5%;   Score 488;  DB 4;  Length 470;
Best Local Similarity   30.0%;   Pred. No. 1.5e-43;
Matches  144;  Conservative  71;  Mismatches  179;  Indels  86;  Gaps  16;

Qy   20  TDASAAASLICGHFAAVDSKFVDTAFAV----DNTYLLFSAYLVFSMQL-GFAMLCAGSV  74
Db   41  TSTSEAPATIAAAPAAAEETPAAPTPTAKLDTGDTSWILISTALVLLMTIPGLALFYGGMV  100

Qy   75  RAKNTMNIMLTNVLDAAAGGLFYYLFGFAPAFGSPSNGFIGKHFFGLK-------DIP  125
Db  101  RKKQNVLSTMFSLSAAILVSLLWVIAGYSIAFSG-----TGAYFGDLSKAMLNGVAFDAL  155

Qy  126  SSSYDYSYF-LYQWAFAIAAAGITSGSIAERTQFVAYLIYSSFLTGFVYPVVSHWFWSPD  184
Db  156  SGTIPESLFVIFQMTFAIITVAILSGSIADRMKYSAFMAFIAIWVLVVYAPITHWVWAAD  215
```

```
Qy  185  GWASAFKITDRLFSTGVIDFPAGSGVVHMVGGIAGLWGALIEGPRMGRFDHAGRAVALRGH  244
Db  216  GW-----LFKAGALDFPAGGTVVHINSGVAGLVAAYMLGKRIG---LGRE-SMAPH  261

Qy  245  SASLVVLGTFLLWFGWYGFNPGSFNKILLTYGNSGNYYGQWSAVGRTAVTTLLAGSTAAL  304
Db  262  NLTLTVIGASLLWVGWFGFNGGS-----------ALGAGARASMAILVTQVAAAAAF  308

Qy  305  TTLFGKRVISGHWNVTDVCNGLLGGFAAITAGCSVVEPWAAIVCGFVASIVLI-ACNKLA  363
Db  309  SWLVVERMIRGKASVLGGASGAVAGLVVITPAAGFVGVGGALVMGLIGGVVCFWGITALK  368

Qy  364  EKVKFDDPLEAAQLHGGCGTWGVIFTALFAKKEYVKEVYGLGRAHGLLMGGGGKLLAAHV  423
Db  369  RLLKADDALDAFGLHAVGGIVGAILTGVFYSDEIIK--------------AANV  408

Qy  424  IQILVIAG----WVS----ATMGPLFWGLNK---LKL-----LRISSEDELAGMDTRHG  467
Db  409  ALAPTFAGQLMVQVEGVLATM--VYSGIATFIILKVFDLIIGLRVNSDDERMGLDLSQHG  466


RESULT 3
US-09-543-681A-6105
; Sequence 6105, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABIL
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 6105
; LENGTH: 433
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-6105

Query Match             17.7%;   Score 468;  DB 4;  Length 433;
Best Local Similarity   28.6%;   Pred. No. 1.9e-41;
Matches  128;  Conservative  83;  Mismatches  188;  Indels  48;  Gaps  11;

Qy   35  AVDSKFVDTAFAVDNTYLLFSAYLVFSMQL-GFAMLCAGSVRAKNTMNIMLTNVLDAAAG  93
Db   23  AIGATTIDIDKA-DNSFMLICSALVFMTIPGIALFYGGLLRSKNVLSLMTQVMLIFSVV  81

Qy   94  GLFYYLFGFAPAFGSPSNGFIGKHFFGLKDIPSSSY-----DYSYFLYQWAFAIAAAGIT  148
Db   82  IILWIMFGYSLAFTAGNKIWGGWSLTFLNNISVDSVAGEINQYVHVVFQGSFAVITVALI  141

Qy  149  SGSIAERTQFVAYLIYSSFLTGFVYPVVSHWFWSPDGWASAPKITDRLFSTGVIDFPAGSG  208
Db  142  VGALGERIRFSALLIFTVIWFTFSYIPIAHMVWGEGGW----------LIDDGALDFPAGGT  192

Qy  209  VVHMVGGIAGLWGALIEGPRMGRFDHAGRAVALRGHSASLVVLGTFLLWFGWYGFNPGSF  268
Db  193  VVHINAAVAALVGAYLLGKR-GDYQH---TAIKPHNLPMVFTGTAVLYIGWFGFNAGSA  247

Qy  269  NKILLTYGNSGNYYGQWSAVGRTA---VTTLLAGSTAALTTLFGKRVISGHWNVTDVCNG  325
Db  248  G---------SANGIAALAFLNTVVATAGAVLAWTFAEWLVRRKPSMLGSCSG  291

Qy  326  LLGGFAAITAGCSVVEPWAAIVCGFVASIV-LIACNKLAKBKVKFDDPLEAAQLHGGCGTW  384
Db  292  CIAGLVAITPAAGTVGPIGALVIGIIAGIIGLWGVVKRWLKADDVCDVFGIHGTCGIA  351

Qy  385  GVIFTALFAKKEYVKEVYGLGRAHGLLMGGGGKLLAAHVIQILVIAGWVSATMGPLFWGL  444
Db  352  GCLLTGIFT----ASFLGGGIGYSEHMTL---GKQVLTQLFSVVITLIWSSVVAYIAFKIA  404
```

```
RESULT 4
US-09-489-039A-12960
; Sequence 12960, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 12960
; LENGTH: 439
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-12960

Query Match       17.1%;   Score 450;   DB 4;   Length 439;
Best Local Similarity   29.3%;   Pred. No. 1.6e-39;
Matches 130;   Conservative  69;   Mismatches 177;   Indels  68;   Gaps  15;

Qy    48  DNTYLLFSAYLVFSMQL-GFAMLCAGSVRAKNTMNIMLTNVLDAAAGGL---FYYLFGFA  103
Db    42  DNAFMMICTALVLFMTIPGIALFYGGLIRGKNVLS-MLTQVI--VTFGLVCVLWVIYGYT  98

Qy   104  FAFGPSNGPIGKHFFG--------LKDIPSSSYDYSFYLYQWAFAIAAAGITSGSI  152
Db    99  LAFGT----GGSFFGSFDWMLKNIELKALMGTFYQYIHVAFQGSFACITVGLIVGAL  152

Qy   153  AERTQFVAYLLYSSFLTGFVYPVVSHWFWSPDGWASAFKITDRLFSTGVIDFAGSGVVHM  212
Db   153  AERIRFSAVLIFVVVWMTLSYPIAHMWG--------GGLLATHGALDFAGGTVVHI  202

Qy   213  VGGIAGLWGALIEGPRMGRFDHAGRAVALRGHSASLVVLGTFLLWFGWYGFNPGSFNKIL  272
Db   203  NAAVAGLVGAYMMGKRVG----FGKE-AFKPHNLPMVFTGTAILYVGWFGFNAGSA----  253

Qy   273  LTYGNSGNYYGQWSAVGRTA----VTTTLAGSTAALTTLFGKRVISGHWNVTDVCNGLLG  328
Db   254  --------SAANEIAALAFVNTVVATAAAILAWTFGEWALRGKPSLLGACSGAIA  300

Qy   329  GFAAITAGCSVVEPWAAIVCGFVASIVLI-ACNKLAEKVKFDDPLEAAQLHGGCGTWGVI  387
Db   301  GLVGVTPACGYIGVGGALIVGIASGLAGIWGVTALKRWLRVDDPCDVFGVHGVCGIVGCI  360

Qy   388  FTALFAKKEYVKEVYGLGRAHGLLMGGGGKLIAAHVIQILVIAGWVSATMGPLFWGLNKL  447
Db   361  LTGIFA---ATSLGGVYAEGVTM-GHQLLVQLESIAITVVWSGVVAFIG--YKVADMT  413

Qy   448  KLLRISSEDELAGMDMTRHGGFAY  471
Db   414  VGLRVPEEQEREGLDVNSHGENAY  437


RESULT 5
US-09-252-991A-23745
; Sequence 23745, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 23745
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23745

Query Match       16.7%;   Score 442;   DB 4;   Length 455;
Best Local Similarity   27.8%;   Pred. No. 1.3e-38;
Matches 125;   Conservative  68;   Mismatches 190;   Indels  66;   Gaps  12;

Qy    48  DNTYLLFSAYLVFSMQL-GFAMLCAGSVRAKNTMNIMLTNVLDAAAGGLFYYLFGFAFAF  106
Db    46  DTAWMLISTALVLLMTIPGLALFYGGMVRAKNVLSIMMQCFAITGLITILWVVYGYSLAF  105

Qy   107  GSPS--------NGFIG--KHFF------GLKDIPSSSYDYSFYLYQWAFAIAAAGITS  149
Db   106  DTTGMEKGVLNFNSFVGGLDKAFLSGLTADGLTSATALFPESVFITFQMTFAIITPALIV  165

Qy   150  GSIAERTQFVAYLLYYSSFLTGFYPVVSHWFWSPDGWASAFKITDRLFSTGVIDFAGSGV  209
Db   166  GAFAERMKFSAMLIFMAVWFTVYAPIAHMVWSGDG--------ALMWDWGVLDFAGGTV  217

Qy   210  VHMVGGIAGLWGALIEGPRMGRFDHAGRAVALRGHSASLVVLGTFLLWFGWYGFNPGSFN  269
Db   218  VHINAGIAGLVACLVLGKRKGY-----PTTPMAPHNLGYTLVGAAMLWIGWFGFNAGS--  270

Qy   270  KILLTYGNSGNYYGQWSAVGRTAVTTTLFGKRVISGHWNVTDVCNGLLGG  329
Db   271  -----AAAAN----GTAGMAMLVTQIATAAAALAWMFAEWITHGKPSALGIASGVVAG  319

Qy   330  FAAITAGCSVVEPWAAIVCGFVASIV-LIACNKLAEKVKFDDPLEAAQLHGGCGTWGVIF  388
Db   320  LVAITPAAGTAGPMGALVIGLASGVICFFAATSLKRALKYDDSLDAFGVHAVGGIVGALL  379

Qy   389  TALFAKKEYVKEVYGLGRAHGLLMGGGGK--LLAAHVIQILVIAGWVSATMGPLFWGLN  445
Db   380  TGIFAAPS------------LGGFGSVEDIGAQFFVQFKGVAFTVVVTFVVTFVILK  424

Qy   446  KLKL--LRISSEDELAGMDMTRHGGFAY  471
Db   425  VLDLVMGLRVTEEEAVGLDLALHNERGY  453


RESULT 6
US-09-328-352-4217
; Sequence 4217, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 4217
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-4217

Query Match       16.7%;   Score 441;   DB 4;   Length 431;
Best Local Similarity   29.6%;   Pred. No. 1.5e-38;
Matches 132;   Conservative  66;   Mismatches 176;   Indels  72;   Gaps  96;

Qy    38  SKFVDTAFA-VDNTYLLFSAYLVFSMQLGFAMLCAGSVRAKNTMNIMLTNVLDAAAGGLF  96
```

```
Db    30  STFFEHAMQNVDLFPLLGAVLVLAMHAGFAFLELGVTVRHKNQVNALSKILTDFALSAIA   89
Qy    97  YYLFGFAFAFGSPSNGFIGKHFFGLKDIPSSSYDYS----YFFLYQWAFAIAAAGITSGSI 152
Db    90  YFFVGYYISY--------GQHFFHEGTVLSNDHGYNLMRCFFLLTFAAAIPA--IISGGI  139
Qy   153  AERTQFVAYLIYSSFLTGFVYPVVSHWFWFSPDGWASAPKITDRLFSTGVIDFAGSGVVHM 212
Db   140  AERAKMRSQAIATLALVALVYPFFEGMVW--NGNYGLQKWLETTFGAAFHDFAGSVVVHA  197
Qy   213  VGGIAGLWGALIEGPRMGRFDHAGRAVALRGHSASLVVLGTFLLMFGWYGFNPGSFNKIL  272
Db   198  MGGWIALAAAVILLGARSGRYKDGRVSAHPPSSIPFLALGSWILIVGWFGFNVMSAQRV-  256
Qy   273  LTYGNSGNYYGQWSAV-GRTAVTTTLA--GSTAALTTLFGKRVISGHWNVTDVCNGLLGG  329
Db   257  ---------DAISGLVAINSLMAMVGGTITANAI------GKNDPGFLHNGPLAG       296
Qy   330  FAATTAGCSVVEPWAAIVCGFVASIVLIACNKLAE-KVKFDDPLEAAQLHGGCGTWGVIF  388
Db   297  LVAICAGSDIVHPVSALVIGGAAGAMFVYLFFYTQNKLKVDDVLGVWPLHGVCGVCGAFGGIA 356
Qy   389  TALFAKK--------EYVKEVGLGRAHGLLMGGGGKLLAAHVIQILVIAGWVSATMGPL   440
Db   357  VGIFGQKWLGGLGGVSFISQLIGTALAIAIALAGG-----------FIVYGILKATIG--  403
Qy   441  FWGLNKLKLLRISSEDELAGMDMTRH                                    466
Db   404  ------------IRLSQEDEFRGADLSIH                                 420
```

RESULT 7
US-09-252-991A-18525
; Sequence 18525, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 18525
; LENGTH: 464
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18525

```
Query Match       15.9%;   Score 420.5;  DB 4;  Length 464;
Best Local Similarity   31.1%;   Pred. No. 2.6e-36;
Matches  137;  Conservative  58;  Mismatches 178;  Indels  67;  Gaps  14;

Qy    49  NT-YLLFSAYLVFSMQLGFAMLCAGSVRAKNTMNIMLTNVLDAAAGGLFYYLFGFAFAFG  107
Db    78  NTLFILLGAVMVLAMHAGFAFLEVGTVRLKNQVNALSKILSDFAISALAYFFVGYWIAYG  137
Qy   108  SPSNGFIGKHFFGLKDIPS-----SSYDYSYFLYQWAFAIAAAGITSGSIAERTQFVAY   161
Db   138  VT-----FFH---PAAALTVDSGYALVKFFLLTFAAAIPAIISGGIAERARFGPQ      185
Qy   162  LIYSSFLTGFVYPVVSHWFWFSPDGWASAPKITDRL---FSTGVIDFAGSGVVHMVGGIAG 218
Db   186  LCATALIVAFVYP----FFEGLVWNGNFGLQEWLKLEFGAPFHDFAGSVVVHALGGWLA   240
Qy   219  LWGALIEGPRMGRFDHAGRAVALRGHSASLVVLGTFLLMFGWYGFNVMSAQTLA        278
Db   241  LAAVLLLGSRNGRY-RDGKLVAMAPSSIPFLALGSWILIIGWFGFNVMSAQTLA------  293
```

```
Qy   279  GNYYGQWSAVGRTAVTTTLA--GSTAALTTLFGKRVISGHWNVTDVCNGLLGGFAAITAG  336
Db   294  ------GVSGLVAVNSLLAMVGGTMA------SLLIGRNDPGFLHNGPLAGLAGLVAVCAG 339
Qy   337  CSVVEPWAAIVCGFVASIVLI-ACNKLAEKVKFDDPLEAAQLHGGCGTWGVIFTALFAKK  395
Db   340  SDLMHPIGALATGLVAGALFVWAFTATQVRWKIDDVLGVWPLHGLCGVWGGIACGIFGQQ  399
Qy   396  EYVKEVGLGRAHGLLMGGGGKLLAAHVIQILVIAGWVSATMGPLFWGLNKLKL-LRISS   454
Db   400  ---------ALGLGGVSLASQALGSL-LGVTVAFAGGLLVYGLMKALLGIRLSQ        444
Qy   455  EDELAGMDMTRHGGFAYAYE                                          474
Db   445  EBEYYGADLSIHKIGAISHE                                          464
```

RESULT 8
US-09-134-001C-3907
; Sequence 3907, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3907
; LENGTH: 437
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3907

```
Query Match       15.4%;   Score 406.5;  DB 4;  Length 437;
Best Local Similarity   26.0%;   Pred. No. 7.7e-35;
Matches  118;  Conservative  80;  Mismatches 191;  Indels  65;  Gaps  13;

Qy    48  DNTYLLFSAYLVFSMQLGFAMLCAGSVRAKNTMNIMLTNVLDAAAGGLFYYLFGFAFAFG  107
Db    26  DTIFLFLCTLLVWLMTPGLSLFYGGLVQSKNALNTVMQSMVAIVTFVWIIIGFSLSFD    85
Qy   108  SPSN-----GFIGKHFFGLK------DIPSSSYDYSYFLYQWAFAIAAAGITSGSIAER   155
Db    86  GGNQWIGGLKFLGLHHVGFETSKTLSPHIPLS---LFMLFQMMFCTIAVSILSGGIAEK   141
Qy   156  TQFVAYLIYSSFLTGFVYPVVSHWFWFSPDGWASAFKITDRLFSTGVIDFAGSGVVHMVGG 215
Db   142  MRFIPYLLFVSLWVLLIYSPVAHWVWG-GGWIS---------KIGAIDYAGGTVVHITSG  191
Qy   216  IAGLWGALIEGPRMGRFDHAGRAVALRGHSASLVLGTFLLMFGWYGFNPGS---FNKIL   272
Db   192  VSGLVLGIMIGIGKCKEKHT-------PHNLLITLIGGILVWLGWYGFNVGSAFTFDHIA  244
Qy   273  LTYGNSGNYYGQWSAVGRTAVTTTLAGSTAALTTLFGKRVISGHWNVTDVCNGLLGGFAA  332
Db   245  MI-----------SFVNTVIGASAGAFGWLIFEYILKKTTSLLGLLSGALSGLVA      288
Qy   333  ITAGCSVVPFWAAIVCGFVASI-VLIACNKLAEKVKFDDPLEAAQLHGGCGTWGVIFTAL  391
Db   289  ITPAAGYVSYMSAMIIAIIGGICCYIVINLLKVKLQYNDALDAFGIHGVGGIIGAVLTGV  348
Qy   392  FAKKEYVKEVGLGRAHGLLMGGGKLLAAHVIQILVIAGWVSATMGPLFWGLNKLKL--   449
Db   349  FQSHQINSAV-----QNGFIYTADFKVV---VIQLGAAIATVFSAIVTFLIARFIKIFT   400
Qy   450  -LRISSEDELAGMDMTRHGGFAYAY-EDDETHKH                            481
```

Db   401 PLATTQEEDKTGLDAIVHGEKAYFYGELNKFNRH 434

RESULT 9
US-09-134-000C-5907
; Sequence 5907, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
;   APPLICANT: Lynn Doucette-Stamm et al
;   TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
;   TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
;   FILE REFERENCE: 032796-032
;   CURRENT APPLICATION NUMBER: US/09/134,000C
;   CURRENT FILING DATE: 1998-08-13
;   PRIOR APPLICATION NUMBER: US 60/055,778
;   PRIOR FILING DATE: 1997-08-15
;   NUMBER OF SEQ ID NOS: 6812
;   SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5907
;   LENGTH: 361
;   TYPE: PRT
;   ORGANISM: Enterococcus faecalis
US-09-134-000C-5907

Query Match   7.7%;   Score 203.5; DB 4; Length 361;
Best Local Similarity 20.5%;   Pred. No. 3.6e-13;
Matches   75; Conservative   72; Mismatches 169; Indels 49; Gaps 10;

Qy   47 VDNTYLLFSAYLVFSMQLGFAMLCAGSVRAKNTMNIMLTNVLDAAAGGLFYLLFGFAFAF 106
             ::| :   : :: :   ::|: | |:|:|
Db   6 MNNLFIFVCFCFMWLMIFGVIIYYVGLVNHRYIHHTLILGLVTIISGTICWLFVGYSLS- 64

Qy   107 GSPSNGFGIGKHFFGLKDIPSSYDYSYFLYQWAFAIAAAGTISGSIAERTQFVAYLIYSS 166
            | : | :: | |:: | :|| :| | :::
Db   65 ------FFGNIQYSIFYSPLASSEIVSILIQLLFCLYSVIMIIGSVLERGNWKYIVLFVP 118

Qy   167 FLTGFVYPVVSHWFWSPDGWASAFKITDRLFSTGVIDRAGSGVVHMVGGIAGLMGALIEG 226
            ||| | : |:|:||| |||:||| |||| | :
Db   119 LWIVFVYAPVCFSLWGHGNW---------LGKIGVLDYSGGLVHTTAGIGSLVLAITSP 169

Qy   227 PRMGRFDHAGRAVALRGHSASLVV-LGTFLLWFGWYGFNPGSFNKILLTYGNSGNYYGQW 285
             :   : || :| | :|| | || | ::: :
Db   170 IRLKN-------SLIFKSQEMIAFVGMLFITLGWFGFNMAPSGKI---------GEE 210

Qy   286 S-AVGRTAVTTTLAGSTAALTTLFGKRVISGHWNVTDVCNGLLGGFAAITAGCSVVEP- 342
            | :||| :   :: ||| |::||| || || || ||| ||| |||| :::
Db   211 SIQIWLNTLISILGGS---ISWPFTQWILIKKVSIYSIMNGIIGGLVGSTCSVGYISPAI 267

Qy   343 --WAAIVCGFVASIVLIACNKLAEKVKFDDPLEAAQLHGGCTWGVIFTALFAKK-EYV 398
             ||| | :: ||| ::   ::|: : : :: : :   :|: :|:|:|| :|:
Db   268 SLLISVIVCTICPIVIHIMHLRIA---NFDDAADSFGMNAVGGIAGSILFGVMAEKGDFF 324

Qy   399 KEVYG 403
            ::|
Db   325 LQLFG 329

RESULT 10
US-09-107-532A-3674
; Sequence 3674, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
;   APPLICANT: Lynn A Doucette-Stamm and David Bush
;   TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
;   TITLE OF INVENTION: ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
;   NUMBER OF SEQUENCES: 7310
;   CORRSPONDENCE ADDRESS:
;   ADDRESSEE: GENOME THERAPEUTICS CORPORATION
;   STREET: 100 Beaver Street
;   CITY: Waltham
;   STATE: Massachusetts
;   COUNTRY: USA
;   ZIP: 02354

COMPUTER READABLE FORM:
  MEDIUM TYPE: CD/ROM ISO9660
  COMPUTER: PC
  OPERATING SYSTEM: <Unknown>
  SOFTWARE: ASCII
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/09/107,532A
  FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: 60/085,598
  FILING DATE: 14 May 1998
  APPLICATION NUMBER: 60/051571
  FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
  NAME: Ariniello, Pamela Deneke
  REGISTRATION NUMBER: 40,489
  REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
  TELEPHONE: (781)893-5007
  TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 3674:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 223 amino acids
    TYPE: amino acid
    TOPOLOGY: linear
  MOLECULE TYPE: protein
  HYPOTHETICAL: YES
  ORIGINAL SOURCE:
    ORGANISM: Enterococcus faecium
  FEATURE:
    NAME/KEY: misc_feature
    LOCATION: (B) LOCATION 1...223
  SEQUENCE DESCRIPTION: SEQ ID NO: 3674:
US-09-107-532A-3674

Query Match   7.3%;   Score 193; DB 4; Length 223;
Best Local Similarity 26.9%;   Pred. No. 2.3e-12;
Matches   60; Conservative   35; Mismatches 102; Indels 26; Gaps 4;

Qy   242 RGHSASLV--VLGTFLLWFGWYGFNPGSFNKILLTYGNSGNYYGQWSAVGRTAVTTTLAG 299
            :||:||: ||:||||||||:|:||||||||| | ::| | : : :|:|
Db   16 RSQHSSLMAAAIGTGLLWFGWFGFNSGG--ALRADEQAVNAFG-------STFIAL 62

Qy   300 STAALTTLFGKRVISGHWNVTDVCNGLLGGFAAITAGCSVVEPWAAIVCGFVASIVLIAC 359
            : :||| | : || ::| :: :: : : ::| |: | : | | :
Db   63 AFAMITWLIIAKVKGNGFDFVDVLTGSVAGLAAITPCAGVVEAKSAMLIGIIAGIVCHAA 122

Qy   360 NKLAEKVKFDDPLEAAQLHGGCGTWGVIFTALFAKKEYVKEVYGLGRAHGLLMGGGGKLL 419
            |||||||||||||||||||||| |:|||||||||||||||||||||||||||||||||||
Db   123 VDFRKKKQWDDALDVWGVHGMGGFGTILIGIFA------------SNSHLLTNPSSWYFL 171

Qy   420 AAHVIQILVIAGWVSATMGPLFWGLNKLKLLRISSEDELAGMD 462
            | |:| | :| | | |:::|:| |||:|:: : :::|
Db   172 GIQVIGVIITAVYAYVLTTVILKSAVHFTTTTTKEEQEEGLD 214

RESULT 11
US-09-107-532A-4102
; Sequence 4102, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
;   APPLICANT: Lynn A Doucette-Stamm and David Bush
;   TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
;   TITLE OF INVENTION: ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
;   NUMBER OF SEQUENCES: 7310
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: GENOME THERAPEUTICS CORPORATION
;   STREET: 100 Beaver Street
;   CITY: Waltham
;   STATE: Massachusetts
;   COUNTRY: USA
;   ZIP: 02354
;   COMPUTER READABLE FORM:

MEDIUM TYPE: CD/ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 4102:
SEQUENCE CHARACTERISTICS:
LENGTH: 224 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...224
SEQUENCE DESCRIPTION: SEQ ID NO: 4102:
US-09-107-532A-4102

Query Match   6.9%; Score 182; DB 4; Length 224;
Best Local Similarity 26.7%; Pred. No. 3.5e-11;
Matches 52; Conservative 21; Mismatches 102; Indels 20; Gaps 4;

Qy   48 DNTYLFSAYLVFSMQLGFAMLCAGSVRAKNTMNIMLTNVLDAAAGGLFYYLFGFAFAFG 107
                :|: ||:||||| ::||| : |||:||::
Db   34 DTAFMLCTAMVCLMTFGLAFFVGGLARKNNILTIMSQSFLSVGITTIMWVFGGFGLAFG 93

Qy 108 SPSNGFIGK--HFFGLKDIPSSSYDY------------SYFLYQWAFAIAAAGITSGSIAERTQ 157
              ::| ||:|: :| :|||| : ::|| |
Db   94 RDIGGVIGNPADFFLMRHVTSFPNMFHGATIPFLMFFLYQLMFIVITVPLMTGAFAGRLN 153

Qy 158 FVAYLIYSSFLTGFVYPVVSHWFWSPDGWASAFKITDRLFSTGVIDFAGSGVVHMVGGIA 217
              :|| | | |||||||| |
Db 154 LKGYIIILVIFWNLLIYFPVCHWI-----WGGGF----LDQMGFRDFAAGAVIHTTAGFG 203

Qy 218 GLWGALIEGPRMGRF 232
             |:|:
Db 204 SLACILTLGQKDSF 218

RESULT 12
US-09-489-039A-9335
; Sequence 9335, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 9335
; LENGTH: 443
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae

US-09-489-039A-9335

Query Match   5.2%; Score 136; DB 4; Length 443;
Best Local Similarity 22.4%; Pred. No. 8.9e-06;
Matches 105; Conservative 55; Mismatches 162; Indels 146; Gaps 25;

Qy   57 YLVFSMQLGFAMLCAGSVRAKNTMNIMLTNV----LDAAAGGLFYYLFGFAFAFGS-PS 110
              ||: :: :|:| ::|
Db   17 YLILLM--IFVVTAVNYGDRATLSIAGTEVAKELGLSAVSMGYIFSAFGWAYLLMQIPG 73

Qy 111 NGFIGKHFFGLKDIPSSSYDYSYFLYQWA----------FAIAAAGITSGSIAERTQFV 159
              || | | ||: |::
Db   74 GWLLDK--FGSKKV------YSYSLFFWSLFTFLQGFIDVFPLAWAGVS---------- 114

Qy 160 AYLIYSSFLTGF----VYP----VVSHWFWSPD-GWASAFKITDRLFS------TGVID 203
             | :   :| |: :|||
Db 115 --MFFMRFMLGFSEAPSFPANARIVAAWFPAKERGTASAIFNAAQYFSLALFSPLLGWLT 172

Qy 204 FA-GSGVVHMVGGIAG-----LWGALI-----EGPRMGR----------FDHAGRAV 239
            | | :   : |||
Db 173 FALGWEHVFTVMGIIGFVLTIIWVKFVHNPTDHPRMSAAELKYISEGGAVVDMDHKKEAT 232

Qy 240 ALRGHSASLV-------VLGTFLLWFGWYGFNPGSFNKILLTYGNSGNYYGQW---SAV 288
                 : || : ||| :
Db 233 PAAGPKMDYIRQLLTNRVMLGVF---FGQYFLN-----TITW------FFLTWFPIYLV 277

Qy 289 GRTAVTTTLAGSTAALTTLFGKRVISGHWNVTDVCNGLLGGFAA---ITAGCSVV-EPWA 344
               : : | |:| ||::
Db 278 QDKGMSILKVGFVASIPALFG-----------FAGGVLGGLFSDYLIGRGCTLTFARKL 325

Qy 345 AIVCGFVASIVLIACNKLAEKVKFPDDPLEAAQLHGGCGTWGVIFTALFAKKSYVK----- 399
                 :| :: :   : :   ::   :
Db 326 PIVLGMLLASSIILCNYTASTPLVITLMALAFFGKGFGALGWPVISDVAPKEIVGLCGGV 385

Qy 400 -EVYG----------------LGRAHGLLMGGGGKLLAAHVIQILVI 429
                                  |:| :: |
Db 386 FNVFGNVASIATPLVIGYIVSELHSFNGALIFVGGSALMMVCYLFVV 433

RESULT 13
US-09-489-039A-9610
; Sequence 9610, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 9610
; LENGTH: 490
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-9610

Query Match   4.6%; Score 121.5; DB 4; Length 490;
Best Local Similarity 19.4%; Pred. No. 0.00037;
Matches 103; Conservative 61; Mismatches 134; Indels 233; Gaps 27;

Qy   13 QLLGPNTTDASAAASLICGHFAAVDSKFVDTAFA--------VDNTYLLF-------SAY 57
              :::   :||
Db   63 QMAGP----ALALVLVCGIF----SFFILRALGELVLHRPSSGSFVSYAREFIGEKAY 114

Qy   58 LVFSMQLGFAMLCAGSVRAKNTMNIMLTNVLDAAAGGLFYYLFGFAFAFGPSNGFIGKH 117
                                             ||| |:: :| |
Db 115 VAGWMYF-------------------VNWAMTGIVDITAVALYMHYWG---AFG------ 146

Qy 118 FFGLKDIPSSSYDYSYFLYQWAFAIAAAGITSG------------SIAERTQFVAY 161
            ||| :: : |||

```
Qy  280  --NYYGQWSAVGRTAV--TTTLAGSTAALTTLFGKRVISGHWNVTDVCNGLLGGFAAITA  335
              ::          ::      :|  ::|       |||| :::
Db  305  VPNYLSQLQGFNARQIGDSLILYGLVGLVQLLLAPLLPRLM--RWLNPKLL--VAGGFAIMAL  360

Qy  336  GC----------SVBPWAAIVCGFVASIVLIACNKLAEKVKFDDPLRAA---------  375
         ||                ::::|
Db  361  GCWMGAHLNADAGRNVIIP-SIVVRGIGQPLIMVALSVLA--VKGLDKAEAGSASALISM  417

Qy  376  --QLHGGCGTWGVIFTALFAKKEVKEVVGLGRAHGLLMG------------GGGKL  418
                                          :|||:|
Db  418  LRNLGGAIGT--ALLTQLVSLRE---------RFHSERIGETLTIFDGALQQRLGGGQV  465
```

RESULT 15
US-09-252-991A-18216
```
; Sequence 18216, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 18216
; LENGTH: 557
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18216
```

```
Query Match           4.5%;   Score 119;  DB 4;  Length 557;
Best Local Similarity 24.0%;  Pred. No. 0.00085;
Matches   85;  Conservative  34;  Mismatches 109;  Indels 126;  Gaps  19;

Qy  164  YSSFLTGFVYPVVSHWFWSPDGWASAFKITDRLFSTGVIDFAGSGVVHMVGGIAGLWGAL  223
         |:||                     ::|::               |||
Db   54  YDFFLYGALAAVISKQFFAGVNDTTAFIFALMAFAAGF------------LVRPFGAL   99

Qy  224  IEGPRMGRFDHAGR------AVALRGHSASLVVLGTFLLWFGWYGFNP-----GSFNKIL  272
         :||||   :|||::          ::                   :|||
Db  100  VFG-RMG--DMIGRKYTFLVTILLMG---------LSTFAV-----GLLPTYASIGVAAPI  144

Qy  273  LT--------------YGNSGNYYGQWSAVGRTAVTTLAGSTAALTLFGKRVISGHWNV  319
         |:                          :
Db  145  LVTLRMLQGLALGGEYGGAAIYVAEHAPANKRGSYTSWIQSTATLGLLLSLLVIL-----  199

Db  320  TDVCNGLLGGFAAITAGCSVVEPWAAIVCGFVASIVLIACN------KLAE----KVKFD  369
                                                            ::
Qy  200  --ACRQLTG------DEFETWGRL-PFLLSIVLLGISTWIRLSMRESPAFLKMKAB  247

Qy  370  DPLEAAQLHGGCGTWG---VIFTALP------AKKEVKEVVGL---------GRAHG  409
                ||||                        :|||
Db  248  GKVSKAPLRESFTQWGNLKVVTALFSINAGQAVITYTAQFVLFFLTQVLKVDGGTANG  307

Qy  410  LLMGGGGKLLAAHVI--QILVIAGWVSATMG---------PLFWGL  444
                  ::|  ::|      :||||          :
Db  308  L-------LIVALVLGAPFFIVAGWLSDRIGRKPVLLAGLLLATLFYPPLFKGL  354
```

Search completed: March 9, 2004, 11:56:43
Job time : 24 secs

```
Db  147  ----DVP--------------QWVFALGALAIVGTMNMIGVKWPAEMEFWFALVKVLAIVAF  190
                               ::|                                 :
Qy  162  LIYSSFLTGFVYPVVSHWFWSPDGWASAFK-ITDR--LFSTGVID----------FAGSGV  209
         ::                                      ::
Db  191  LVVGTIFLGSGKPL--------DGNATGFHLJITDNGGFFPHGLLPALVIVQGVVFAFAS-  241

Qy  210  VHMVGGIAG---------------LWGALIEGPRMGRFDHAGRAVALRGHSASLVVLGT  253
         : ::                          :                :|||
Db  242  IELVGTAAGECKDPETMVPKAINSVIW------RIGLF----------YVGSVVLLVL  283

Qy  254  FLLWFGWY-GFNP----------GS-FNKILLTYG----NSGNYYGQWSAVGRTAVT  294
         | :                     :::|
Db  284  LLPWNAYQAGQSPFVTFFSKLGVPYIGSVMNIVVLTRALSSLNSGLY-----STGRILRS  338

Qy  295  TTTLAGSTAALTTLFGKRVISGHWNVTDVCNGLLGGFAAITAGCSVVEPWAAIVCGFVASI  354
         :::                                          :
Db  339  MSMGGSAPKFMSKMSRHHV--------------------------PYAGILATLGVTV  370

Qy  355  VLIACNKLAEKVKFDDPLEAAQLHGGCGTWGVIFTALFAKKEVKEVVGLGRA-------  407
          :|                    :|||  |                  ::
Db  371  VGVFLNYLVPSQVFEIVLNVASL-GIIASWGFIVVCQMRLRKAIKE----GKAAKVSFRM  425

Qy  408  --------------------HGLLMGGGGKLLAAHVIQILVIAGW  432
                             :|||    :||  :::|:||:|||
Db  426  PGAPFTSWLTLLFLFSVLVLMAFDYPNGTYTIGSIPLLA------VLLVAGW  471
```

RESULT 14
US-09-252-991A-19322
```
; Sequence 19322, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19322
; LENGTH: 517
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19322
```

```
Query Match           4.5%;   Score 119.5;  DB 4;  Length 517;
Best Local Similarity 22.3%;  Pred. No. 0.00067;
Matches  107;  Conservative  54;  Mismatches 193;  Indels 125;  Gaps  24;

Qy   14  LLGPNTDASAAASLICGHFAAV--DSKFVDTAFAVDNTYLL-FSAYL--VFSMQ------   63
         :                       :|::               |||:
Db   38  MAGMNVHVTSAALPEIEGALGATPEEGSWISTAYLVAEISMIPLTAWLVEVFSLRRVMLL   97

Qy   64  --LGFAMLCAGSVRAKNTMNIMLTNVLDAAAGGLFYYLFGFAFAFGSPSNGFIGKHFFGL  121
           ::|                        :   ||||
Db   98  GSLVFLLSSLSCALAPNLSTLILIRVIQGASGAVLIPL-----------SMQLI------L  141

Qy  122  KDIPSSSYDYSYFLYQWAFAIA-AAGITSGSIAERTQFVAYLIYSSFLTGFVYPVVSHWF  180
             |||                         ::|           :::
Db  142  TELPSSSRIPLGMALIFSLSNSVAQRAGPSIGGWLADAYSWRWIFLIQLLPGIALLAAVAWS  201

Qy  181  WSP-DGWASAFKITDRLFSTGVIDFAGSGVVHMVGGIAGLWGALIEGPRMGRFDHAG-RA  238
          ||                                            :|||||
Db  202  IRPRDGDRERLRQADWL---------GIGAMVAGLGALQIVLEGGRRDWFESCGFIRT  250

Qy  239  VALRGHSASLVVLGTFL----LW-------FGWYGFNPGSFNKILLTYGNSG----  279
         :|||                 |                |||
Db  251  FAV------LAVLALLLFVQRQLWGARPFINLRLLGSYNFGVSSLAMAVFGAATFGLVFL  304
```

```
                    GenCore version 5.1.6
              Copyright (c) 1993 - 2004 Compugen Ltd.

CM protein - protein search, using sw model

Run on:        March  9, 2004, 11:55:40 ; Search time 39 Seconds
                                           (without alignments)
                                           2707.093 Million cell updates/sec

Title:          US-10-033-109-4
Perfect score:  2639
Sequence:       1 MSLPACPABQLAQLLGPNTT.........HGMQLRRVGPNASSTPTTDE 500

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       809742 seqs, 211153259 residues

Total number of hits satisfying chosen parameters:          809742

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :      Published Applications AA:*
         1: /cgn2_6/ptcdata/1/pubpaa/US07_PUBCOMB.pep:*
         2: /cgn2_6/ptcdata/1/pubpaa/PCT_NEW_PUB.pep:*
         3: /cgn2_6/ptcdata/1/pubpaa/US06_NEW_PUB.pep:*
         4: /cgn2_6/ptcdata/1/pubpaa/US06_PUBCOMB.pep:*
         5: /cgn2_6/ptcdata/1/pubpaa/US07_NEW_PUB.pep:*
         6: /cgn2_6/ptcdata/1/pubpaa/PCTUS_PUBCOMB.pep:*
         7: /cgn2_6/ptcdata/1/pubpaa/US08_NEW_PUB.pep:*
         8: /cgn2_6/ptcdata/1/pubpaa/US08_PUBCOMB.pep:*
         9: /cgn2_6/ptcdata/1/pubpaa/US09A_PUBCOMB.pep:*
        10: /cgn2_6/ptcdata/1/pubpaa/US09B_PUBCOMB.pep:*
        11: /cgn2_6/ptcdata/1/pubpaa/US09C_PUBCOMB.pep:*
        12: /cgn2_6/ptcdata/1/pubpaa/US09_NEW_PUB.pep:*
        13: /cgn2_6/ptcdata/1/pubpaa/US10A_PUBCOMB.pep:*
        14: /cgn2_6/ptcdata/1/pubpaa/US10B_PUBCOMB.pep:*
        15: /cgn2_6/ptcdata/1/pubpaa/US10C_PUBCOMB.pep:*
        16: /cgn2_6/ptcdata/1/pubpaa/US10_NEW_PUB.pep:*
        17: /cgn2_6/ptcdata/1/pubpaa/US60_NEW_PUB.pep:*
        18: /cgn2_6/ptcdata/1/pubpaa/US60_PUBCOMB.pep:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Match Length	DB	ID	Description
1	2639	100.0	500	13	US-10-033-109-4	Sequence 4, Appli
2	2129.5	80.7	501	8	US-08-635-967-2	Sequence 2, Appli
3	1954	74.0	494	13	US-10-033-109-6	Sequence 6, Appli
4	775	29.4	459	9	US-09-795-693-32	Sequence 32, Appli
5	775	29.4	459	14	US-10-156-239-32	Sequence 32, Appl
6	775	29.4	459	14	US-10-199-485-32	Sequence 100, Appl
7	707.5	26.8	395	15	US-10-190-115-100	Sequence 100, App
8	707.5	26.8	395	15	US-10-369-072-100	Sequence 250, App
9	677.5	25.7	373	15	US-10-074-978A-250	Sequence 2, Appli
10	613	23.2	183	13	US-10-033-109-2	Sequence 13126, A
11	535	20.3	436	14	US-10-156-761-13126	Sequence 10188, A
12	483	18.3	446	14	US-10-156-761-10188	Sequence 192, App
13	449.5	17.0	492	9	US-09-801-368-192	Sequence 6981, Ap
14	430.5	16.3	438	9	US-09-738-626-6981	Sequence 194, App
15	414	15.7	499	9	US-09-801-368-194	

16	363.5	13.8	452	9	US-09-738-626-6968	Sequence 6968, Ap
17	335.5	12.7	470	13	US-10-033-109-14	Sequence 14, Appl
18	327.5	12.4	486	13	US-10-033-109-12	Sequence 12, Appl
19	314	11.9	497	13	US-10-033-109-10	Sequence 10, Appl
20	180	6.8	458	15	US-10-190-115-59	Sequence 59, Appl
21	180	6.8	458	15	US-10-369-072-59	Sequence 59, Appl
22	173.5	6.6	458	15	US-10-190-115-55	Sequence 55, Appl
23	173.5	6.6	458	15	US-10-190-115-56	Sequence 56, Appl
24	173.5	6.6	458	15	US-10-369-072-55	Sequence 55, Appl
25	173.5	6.6	458	15	US-10-369-072-56	Sequence 56, Appl
26	165	6.3	488	15	US-10-074-978A-248	Sequence 248, App
27	158	6.0	458	15	US-10-190-115-57	Sequence 57, Appl
28	158	6.0	458	15	US-10-369-072-57	Sequence 57, Appl
29	157	5.9	467	15	US-10-074-978A-247	Sequence 247, App
30	151	5.7	498	9	US-09-949-145-4	Sequence 4, Appli
31	151	5.7	498	15	US-10-074-978A-245	Sequence 245, App
32	148	5.6	455	15	US-10-190-115-58	Sequence 58, Appl
33	148	5.6	455	15	US-10-369-072-58	Sequence 58, Appl
34	147	5.6	459	15	US-10-074-978A-246	Sequence 246, App
35	139	5.3	479	9	US-09-949-145-3	Sequence 3, Appli
36	139	5.3	479	15	US-10-074-978A-244	Sequence 244, App
37	134	5.1	431	15	US-10-369-493-13974	Sequence 13974, A
38	132	5.0	432	15	US-10-369-493-15510	Sequence 15510, A
39	132	5.0	432	15	US-10-369-493-15882	Sequence 15882, A
40	132	5.0	432	15	US-10-369-493-16254	Sequence 16254, A
41	131.5	5.0	388	15	US-10-074-978A-249	Sequence 249, App
42	131.5	5.0	445	15	US-10-074-978A-42	Sequence 42, Appl
43	131.5	5.0	448	15	US-10-190-115-22	Sequence 22, Appl
44	131.5	5.0	448	15	US-10-369-072-22	Sequence 22, Appl
45	130.5	4.9	326	15	US-10-264-237-2412	Sequence 2412, Ap

ALIGNMENTS

```
RESULT 1
US-10-033-109-4
; Sequence 4, Application US/10033109
; Publication No. US2002014239OA1
; GENERAL INFORMATION:
;   APPLICANT: Allen, Stephen M.
;   APPLICANT: Rafalski, J. Antoni
;   APPLICANT: Sakai, Hajime
;   TITLE OF INVENTION: Nitrogen Transport Metabolism
;   FILE REFERENCE: BB-1210
;   CURRENT APPLICATION NUMBER: US/10/033,109
;   CURRENT FILING DATE: 2001-12-28
;   PRIOR APPLICATION NUMBER: EARLIER FILING DATE: 1999-08-27
;   PRIOR APPLICATION NUMBER: EARLIER FILING DATE: 09/384,625
;   PRIOR APPLICATION NUMBER: EARLIER FILING DATE: 60/098,248
;   PRIOR FILING DATE: EARLIER FILING DATE: 28 August 1998
;   NUMBER OF SEQ ID NOS: 14
;   SOFTWARE: Microsoft Office 97
; SEQ ID NO 4
;   LENGTH: 500
;   TYPE: PRT
;   ORGANISM: Glycine max
US-10-033-109-4

  Query Match             100.0%;  Score 2639;  DB 13;  Length 500;
  Best Local Similarity   100.0%;  Pred. No. 6.7e-256;
  Matches  500;  Conservative   0;  Mismatches   0;  Indels    0;  Gaps    0;

Qy         1 MSLPACPABQLAQLLGPNTTDASAAASLICGHFAAVDSKFVDTAFAVDNTYLLFSAYLVF 60
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1 MSLPACPAEQLAQLLGPNTTDASAAASLICGHFAAVDSKFVDTAFAVDNTYLLFSAYLVF 60

Qy        61 SMQLGFAMLCAGSVRAKNTNIMLJNVLDAAAGGLFYLFGFAFAFGSPSNGFIGKHFFG 120
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        61 SMQLGFAMLCAGSVRAKNTNIMLJNVLDAAAGGLFYLFGFAFAFGSPSNGFIGKHFFG 120

Qy       121 LKDIPSSSYDYSYFLYQWAPAIAAAGITSGSIAERTQFVAYLIYSSFLTGFVYPVVSHWF 180
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```

```
Db   121 LKDIPSSSYDYSYFLYQWAFAIAAAGITSGSIAERTQFVAYLIYSSFLTGFVYPVVSHWF 180

Qy   181 WSPDGWASAFKITDRLFSTGVIDFAGSGVVHMVGGIAGLWGALIEGPRMGRFDHAGRAVA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 WSPDGWASAFKITDRLFSTGVIDFAGSGVVHMVGGIAGLWGALIEGPRMGRFDHAGRAVA 240

Qy   241 LRGHSASLVVLGTFLLWFGWYGFNPGSFNKILLTYGNSGNYYGQWSAVGRTAVTTTLAGS 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 LRGHSASLVVLGTFLLWFGWYGFNPGSFNKILLTYGNSGNYYGQWSAVGRTAVTTTLAGS 300

Qy   301 TAALTTLFGKRVISGHWNVTDVCNGLLGGFAAITAGCSVVEPWAAIVCGFVASIVLIACN 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 TAALTTLFGKRVISGHWNVTDVCNGLLGGFAAITAGCSVVEPWAAIVCGFVASIVLIACN 360

Qy   361 KLAEKVKFDDPLEAAQLHGGCGTWGVIFTALFAKKEYVKEVYGLGRAHGLLMGGGGKLLA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 KLAEKVKFDDPLEAAQLHGGCGTWGVIFTALFAKKEYVKEVYGLGRAHGLLMGGGGKLLA 420

Qy   421 AHVIQILVIAGWVSATMGPLFWGLNKLKLLRISSEDELAGMDMTRHGGFAYAYEDDETHK 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 AHVIQILVIAGWVSATMGPLFWGLNKLKLLRISSEDELAGMDMTRHGGFAYAYEDDETHK 480

Qy   481 HGMQLRRVGPNASSTPTTDE 500
              ||||||||||||||||||||
Db   481 HGMQLRRVGPNASSTPTTDE 500

RESULT 2
US-08-635-967-2
; Sequence 2, Application US/08635967
; Publication No. US20010003848A1
; GENERAL INFORMATION:
;   APPLICANT: FROMMER, Wolf-Bernd
;   APPLICANT: NINNEMAN, Olaf
;   TITLE OF INVENTION: DNA SEQUENCES FOR AMMONIUM TRANSPORTER,
;   TITLE OF INVENTION: PLASMIDS, BACTERIA, YEASTS, PLANT CELLS AND PLANTS
;   TITLE OF INVENTION: CONTAINING THE TRANSPORTER
;   NUMBER OF SEQUENCES: 2
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Ostrolenk, Faber, Gerb & Soffen, LLP
;   STREET: 1180 Avenue of the Americas
;   CITY: New York
;   STATE: NY
;   COUNTRY: US
;   ZIP: 10036-8403
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/635,967
;   FILING DATE: 29-APR-1996
;   CLASSIFICATION: 800
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: PCT/EP 94/03499
;   FILING DATE: 24-OCT-1994
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: DE P 43 37 597.9
;   FILING DATE: 28-OCT-1993
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Meilman, Edward
;   REGISTRATION NUMBER: 24,735
;   REFERENCE/DOCKET NUMBER: P/951-120
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (212) 382-0700
;   TELEFAX: (212) 382-0888
;   TELEX: 236925
; INFORMATION FOR SEQ ID NO: 2:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 501 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-08-635-967-2

  Query Match   80.7%; Score 2129.5; DB 8; Length 501;
  Best Local Similarity   80.9%; Pred. No. 9.1e-205;
  Matches 407; Conservative   42; Mismatches   39; Indels   15; Gaps   7;

Qy   5 ACPAEQLAQLLGPNTTDASAAASLICGHFAAVDSKFVDTAFAVDNTYLLFSAYLVFSMQL 64
              :| | || ||||| |:|||: ||| |::||:|||||:|||||||||||||||||
Db   2 SCSATDLAVLLGPN---ATAAANYICGQLGDVNNKFIDTAFAIDNTYLLFSAYLVFSMQL 58

Qy   65 GFAMLCAGSVRAKNTMNIMLTNVLDAAAGGLFYYLFGFAFAFGSPSNGFIGKHFFGLKDI 124
              |||||||||||||||||||||||||||||||||||||:|||||||||||||||:||||||
Db   59 GFAMLCAGSVRAKNTMNIMLTNVLDAAAGGLFYYLFGYAFAFGSPSNGFIGKHYFGLKDI 118

Qy   125 PSSSYDYSYFLYQWAFAIAAAGITSGSIAERTQFVAYLIYSSFLTGFVYPVVSHWFWSPD 184
              |::| ||| ||||||||||||||||||||||||||||||||||||||||||||||||| |
Db   119 PTASADYSNFLYQWAFAIAAAGITSGSIAERTQFVAYLIYSSFLTGFVYPVVSHWFWSVD 178

Qy   185 GWASAFKIT-DRLFSTGVIDFAGSGVVHMVGGIAGLWGALIEGPRMGRFDHAGRAVALRG 243
              |||| |: | ||||| |||||||||||||||||||||||||||:||||: |||:||||
Db   179 GWASPFRTDGDLLFSTGAIDFAGSGVVHMVGGIAGLWGALIEGPRLGRFDNGGRAIALRG 238

Qy   244 HSASLVVLGTFLLWFGWYGFNPGSFNKILLTYGNSGNYYGQWSAVGRTAVTTTLAGSTAA 303
              |||||||||||||||||||||||||||||:|| :| | ||||||||||||||||| |||
Db   239 HSASLVVLGTFLLWFGWYGFNPGSFNKILVTY-ETGTYNGQWSAVGRTAVTTTLAGCTAA 297

Qy   304 LTTLFGKRVISGHWNVTDVCNGLLGGFAAITAGCSVVEPWAAIVCGFVASIVLIACNKLA 363
              ||||||||::||||||||||||||||||||||| |||||||||||:|||||::||: |||||
Db   298 LTTLFGKRLLSGHWNVTDVCNGLLGGFAAITGGCSVVEPWAAIICGFVAALVLLGCNKLA 357

Qy   364 EKVKFDDPLEAAQLHGGCGTWGVIFTALFAKKEYVKEVYG--LGRAHGLLMGGGGKLLAA 421
              ||:|:|||||||||||||| ||:|||||||:::|: ::|| || ||| |||||||| |
Db   358 EKLKYDDPLEAAQLHGGCGAWGLIFTALFAQEKYLNQIYGNKPGRPHGLFMGGGGKLLGA 417

Qy   422 HVIQILVIAGWVSATMGPLFWGLNKLKLLRISSEDELAGMDMTRHGGFAYAY-EDDETHK 480
               :|||:|| |||||||| ||: | |:||||||||||:||||||||||||| | :|||:||
Db   418 QLIQIIVITGWVSATMGTLFFILKKMKLLRISSEDEMAGMDMTRHGGFAYMYFDDDESHK 477

Qy   481 HGMQLRRVGP------NASSTPT 497
                :||||| | |::|||
Db   478 -AIQLRRVEPRSPSPSGANTTPT 499

RESULT 3
US-10-033-109-6
; Sequence 6, Application US/10033109
; Publication No. US20020142390A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen M.
; APPLICANT: Rafalski, J. Antoni
; APPLICANT: Sakai, Hajime
; TITLE OF INVENTION: Nitrogen Transport Metabolism
; FILE REFERENCE: BB-1210
; CURRENT APPLICATION NUMBER: US/10/033,109
; CURRENT FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/384,625
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/098,248
; PRIOR FILING DATE: EARLIER FILING DATE: 28 August 1998
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 6
;   LENGTH: 494
;   TYPE: PRT
;   ORGANISM: Triticum aestivum
US-10-033-109-6

  Query Match   74.0%; Score 1954; DB 13; Length 494;
  Best Local Similarity   74.8%; Pred. No. 3.6e-187;
  Matches 374; Conservative   46; Mismatches   62; Indels   18; Gaps   5;

Qy   3 LPACPAEQLAQLLGPNTTDASAAASLICGHFAAVDSKFVDTAFAVDNTYLLFSAYLVFSM 62
```

```

```
Db 121 GAVAERIKFSAYLLFSALLGTLVYPPVAHWVWGELIVGGWLAKLGVLIVLILKTAIDFAGS 180
Qy 208 GVVHMVGGIAGLMGALIEGPRMGRF--DHAGRAVALRGHSASLVVLGTPLLWFGWYGFNP 265
Db 181 TVVHIVGGVAGLAAALVLGPRIGRFPDDETGKPEAIRPHNLPFAVLGTFPLLWFGWFGFNA 240
Qy 266 GSFNKILLTYGNSGNYYGQWSAVGRTAVTTLAGSTAALTTLFGKRVISGHWNVTDVCNG 325
Db 241 GS--ALTANGRAAAIGAGWSTVARAAVNTNLAAAAGALTWLLISRLKTGKPTVLGLANG 297
Qy 326 LLGGFAAI-TAGCSVVEPWAAIVCGFVASIVLI-----ACNKLAEKVKFDDPLEAAQLHG 379
Db 298 ALAGLVAIGTPACGVVSPWGALIIGLVAGVLSVLGVKYLTPKLKEKLGIDDPLDVFPVHG 357
Qy 380 GCGTWGVIFTALFAKKEY----VKEVYGL--GRAHGLLMGGGG-KLLAAHVIQILVIAG 431
Db 358 VGGIWGGIAVGIFAAPKVNNIGFPEEYGASTSGISGGLLYGNGGFKQLGVQLGIGIAVILA 417
Qy 432 W---VSATMGPLFWGLNKLKLLRISSEDELAGMDTRHGGFPAY 471
Db 418 YAFGVTFILAKLL-GLTLGGKLRVSEEBEKVGLDLAEHGETAY 459
```

RESULT 5
US-10-156-239-32
; Sequence 32, Application US/10156239
; Publication No. US2003003 6074A1
; GENERAL INFORMATION:
;  APPLICANT: Glucksmann, Maria A.
;  APPLICANT: Kapeller-Libermann, Rosana
;  TITLE OF INVENTION: No. US20030036074A1e1 Nucleic Acid Sequences Encoding Human Trans
;  TITLE OF INVENTION: ATPase Molecule, A Human Ubiquitin Hydrolase-Like Molecule, A Hu
;  TITLE OF INVENTION: Ubiquitin Conjugating Enzyme-Like Molecule, and Uses Therefor
;  FILE REFERENCE: 35800/247645
;  CURRENT APPLICATION NUMBER: US/10/156,239
;  CURRENT FILING DATE: 2002-05-24
;  PRIOR APPLICATION NUMBER: 09/795,693
;  PRIOR APPLICATION FILING DATE: 2001-02-28
;  PRIOR APPLICATION NUMBER: 60/185,906
;  PRIOR FILING DATE: 2000-02-29
;  PRIOR APPLICATION NUMBER: 09/809,557
;  PRIOR FILING DATE: 2001-03-15
;  PRIOR APPLICATION NUMBER: 60/192,018
;  PRIOR FILING DATE: 2000-03-24
;  PRIOR APPLICATION NUMBER: 09/808,568
;  PRIOR FILING DATE: 2001-03-14
;  PRIOR APPLICATION NUMBER: 60/191,790
;  PRIOR FILING DATE: 2000-03-24
;  PRIOR APPLICATION NUMBER: 09/808,767
;  PRIOR FILING DATE: 2001-03-15
;  PRIOR APPLICATION NUMBER: 60/191,781
;  PRIOR FILING DATE: 2000-03-24
;  NUMBER OF SEQ ID NOS: 60
;  SOFTWARE: FastSEQ for Windows Version 4.0
;  SEQ ID NO 32
;  LENGTH: 459
;  TYPE: PRT
;  ORGANISM: Artificial Sequence
;  FEATURE:
;  OTHER INFORMATION: Pfam consensus sequence
US-10-156-239-32

Query Match     29.4%;  Score 775;  DB 14;  Length 459;
Best Local Similarity   41.3%;  Pred. No. 6.8e-69;
Matches 191;  Conservative   60;  Mismatches 162;  Indels  50;  Gaps  17;

```
Qy   52 LLFSAYLVFSMQLGFAMLCAGSVRAKNTMN-IMLTNVLDAAAG--GLFYYLFGFAFAF-- 106
        |: :||:: :||  ||||:| |||:||: : :| |||::: :|| : :||||||||:
Db    4 LISAALVIFMQPGFALLESGLVRSKNVLNFILMKNFVDLAIGICVLAYVLFGYSLAFGD 63

Qy  107 --GSPSNGFIGKHFF---------GLKD--IPSSSYDYSFYLYQWAFAI-AAAGITS 149
          :|||||||||              |||:  |::|| :|||| :| :||:|||||
Db   64 SYGEPGNGFIGNGLVWLFKFLGVSAAGIQDGTLPDG--LPFFLFQLMFAAKTAATIS 120
```

```
Db    1 MSATCAADLGPLLG--AAAANATDYLCNRFA------DTTSAVDSTYLLFSAYLVPAM 50
Qy   63 QLGFAMLCAGSVRAKNTMNIMLTNVLDAAAGLFYYLFGFAFAFGSPSNGFIGKHFFGLK 122
Db   51 QLGFAMLCAGSVRAKNTMNIMLTNVLDAAAGALFYYLFGFAFAFGTPSNGFIGKHFFGLK 110
Qy  123 DIPSSSYDYSFLYQWAFAIAAAGITSGSIAERTQFVAYLIYSSFLTGFVYPVVSHWFWS 182
Db  111 DMPQTGFDYSFFLFQWAFAIAAAGITSGSIAERTQFVAYLIYSAFLTGFVYPVVSHWIWS 170
Qy  183 PDGWASAFKIT-DRLFSTGVIDFAGSGVVHMVGGIAGLWGALIEGPRMGRFDHAGRAVAL 241
Db  171 VDGWASAARTSGPLLFKSGVIDFAGSGVVHMVGGIAGFWGALIEGPRIGRFDHAGRSVAL 230
Qy  242 RGHSASLVVLGTFLLWFGWYGFNPGSFNKILLTYGNSGNYYGQWSAVGRTAVTTTLAGST 301
Db  231 KGHSASLVVLGTFLLWFGWYGFNPGSFVTILKSYGPPGSINGQWSGVGRTAVTTTTLAGSV 290
Qy  302 AALTTLFGKRVISGHWNVTDVCNGLLGGFAAITAGCSVVEPWAAIVCGFVASIVLIACNK 361
Db  291 AALTTLFGKRLQTGHWNVVDVCNGLLGGFAAITAGCSSVDPWAAVICGFVSAWVLIGLNA 350
Qy  362 LAEKVKFDDPLEAAQLHGGCGTWGVIFTALFAKKEYVKEVYGLGRAHGLLMGGGKLLAA 421
Db  351 LAGRLKYDDDPLEAAQLHGGCGAWGIIFTALFAKKQYVEEIYGAGRPYGLFLGGGRLLAA 410
Qy  422 HVIQILVIAGWVSATMGPLFWGLNKLKLLRISSEDELAGMDTRHGGFAYAYEDDETHKH 481
Db  411 HIVQIIVIAGFVSCTMGPLFLALKKLGLLRISAEDEMAGMDLTRHGGFAYVYHDDDEHDK 470
Qy  482 --GMQLR----RVGFNASS 494
Db  471 SVGGFMLRSAQTRVEPAAAA 490
```

RESULT 4
US-09-795-693-32
; Sequence 32, Application US/09795693
; Patent No. US20020068710A1
; GENERAL INFORMATION:
; APPLICANT: Glucksmann, Maria A.
; TITLE OF INVENTION: 20685, 579, 17114, 23821, 33894, and
; TITLE OF INVENTION: 32613, No. US2002006871OA1e1 Human Transporters
; FILE REFERENCE: 35800/209292
; CURRENT APPLICATION NUMBER: US/09/795,693
; CURRENT FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/185,906
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 32
; LENGTH: 459
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Pfam consensus sequence
US-09-795-693-32

Query Match   29.4%; Score 775; DB 9; Length 459;
Best Local Similarity   41.3%; Pred. No. 6.8e-69;
Matches 191; Conservative   60; Mismatches 162; Indels 50; Gaps 17;

```
Qy   52 LLFSAYLVFSMQLGFAMLCAGSVRAKNTMN-IMLTNVLDAAAG--GLFYYLFGFAFAF-- 106
        |: :||:: :||  ||||:| |||:||: : :| |||::: :|| : :||||||||:
Db    4 LISAALVIFMQPGFALLESGLVRSKNVLNFILMKNFVDLAIGICVLAYVLFGYSLAFGD 63

Qy  107 --GSPSNGFIGKHFF---------GLKD--IPSSSYDYSFYLYQWAFAI-AAAGITS 149
          :|||||||||              |||:  |::|| :|||| :| :||:|||||
Db   64 SYGEPGNGFIGNGLVWLFKFLGVSAAGIQDGTLPDG--LPFFLFQLMFAAKTAATIS 120

Qy  150 GSIAERTQFVAYLIYSSFLTGFVYPVVSHWFWSP--DGWASAFKITDRLFSTGVIDFAGS 207
        || |||||| ||||:|| :  ||| : : :|||:|     || :||: ||:|:||||:||
Db  150 GSIAERTQFVAYLIYSSFLTGFVYPVVSHWFWSP--DGWASAFKITDRLFSTGVIDFAGS 207
```

```
Db   241  GS---ALTANGRAAAIGAGWSTVARAAVNTNLAAAAGALTWLLISRLKTGKPTVLGLANG  297
Qy   326  LLGGFAAI-TAGCSVVEPWAAIVCGFVASIVLI-----ACNKLAEKVKFDDPLEAAQLHG  379
Db   298  ALAGLVAIGTPACGVVSPWGALIIGLVAGVLSVLGVKYLTPKLKEKLGIDDPLDVFPVHG  357
Qy   380  GCGTWGVIFTALFAKKEY----VKEVYGL--GRAHGLLMGGGG-KLLAAHVIQILVIAG   431
Db   358  VGGIWGGIAVGIFAAPKVNNIGFPEEYGASTSGISGGLLYGNGGFKQLGVQLIGIAVILA  417
Qy   432  W---VSATMGPLFWGLNKLKLLRISSEDELAGMDTRHGGFAY  471
Db   418  YAFGVTFILAKLL-GLTLGGKLRVSEEEKVGLDLAEHGETAY  459
```

RESULT 7
US-10-190-115-100
; Sequence 100, Application US/10190115
; Publication No. US20030207394A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook, John P. II
; APPLICANT: Boldog, Ferenc L.
; APPLICANT: Burgess, Catherine E.
; APPLICANT: Casman, Stacie J.
; APPLICANT: Grosse, William M.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Ji, Weizhen
; APPLICANT: Lepley, Denise M.
; APPLICANT: Liu, Xiaohong
; APPLICANT: Mezick, Amanda J.
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Patturajan, Meera
; APPLICANT: Rastelli, Luca
; APPLICANT: Shen, Lei
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Spaderna, Steven K.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Szekeres, Edward S. Jr.
; APPLICANT: Taupier, Raymond J. Jr.
; APPLICANT: Tchernev, Velizar T.
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Voss, Edward Z.
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-050 CIP
; CURRENT APPLICATION NUMBER: US/10/190,115
; CURRENT FILING DATE: 2003-02-10
; PRIOR APPLICATION NUMBER: 60/303,168
; PRIOR FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 60/368,996
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 60/386,816
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: 60/215,854
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/215,856
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/215,902
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/216,585,
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/216,586
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/216,722
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/218,622
; PRIOR FILING DATE: 2000-07-17
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 136
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 100
; LENGTH: 395

```
Qy   150  GSIAERTQFVAYLIYSSFLTGFVYPVVSHWFWSP--DGWASAFKITDRLFSTGVIDFAGS  207
Db   121  GAVAERIKFSAYLLFSALLGTLVYPPVAHWWGELVGGWLAKLGVLVLILKTKAIDFAGS   180
Qy   208  GVHMVGGIAGLWGALIEGPRMGRF--DHAGRAVALRGHSASLVVLGTFLLWFGWYGFNP   265
Db   181  TVVHIVGGVAGLAAALVLGPRIGRFPDDETGKPEAIRPHNLPFAVLGTFLLWFGWFGFNA  240
Qy   266  GSFNKILLTYGNSGNYYGQWSAVGRTAVTTLAGSTAALTTLFGKRVISGHWNVTDVCNG  325
Db   241  GS---ALTANGRAAAIGAGWSTVARAAVNTNLAAAAGALTWLLISRLKTGKPTVLGLANG  297
Qy   326  LLGGFAAI-TAGCSVVEPWAAIVCGFVASIVLI-----ACNKLAEKVKFDDPLEAAQLHG  379
Db   298  ALAGLVAIGTPACGVVSPWGALIIGLVAGVLSVLGVKYLTPKLKEKLGIDDPLDVFPVHG  357
Qy   380  GCGTWGVIFTALFAKKEY----VKEVYGL--GRAHGLLMGGGG-KLLAAHVIQILVIAG   431
Db   358  VGGIWGGIAVGIFAAPKVNNIGFPEEYGASTSGISGGLLYGNGGFKQLGVQLIGIAVILA  417
Qy   432  W---VSATMGPLFWGLNKLKLLRISSEDELAGMDTRHGGFAY  471
Db   418  YAFGVTFILAKLL-GLTLGGKLRVSEEEKVGLDLAEHGETAY  459
```

RESULT 6
US-10-199-485-32
; Sequence 32, Application US/10199485
; Publication No. US20030077626A1
; GENERAL INFORMATION:
; APPLICANT: Glucksmann, Maria A.
; APPLICANT: Silos-Santiago, Inmaculada
; TITLE OF INVENTION: 20685, 579, 17114, 23821, 33894, and
; TITLE OF INVENTION: 32613, No. US2003007762Ale1 Human Transporters
; FILE REFERENCE: 35800/249468
; CURRENT APPLICATION NUMBER: US/10/199,485
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 09/795,693
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/185,906
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 32
; LENGTH: 459
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Pfam consensus sequence
US-10-199-485-32

Query Match   29.4%; Score 775; DB 14; Length 459;
Best Local Similarity 41.3%; Pred. No. 6.8e-69;
Matches 191; Conservative 60; Mismatches 162; Indels 50; Gaps 17;

```
Qy    52  LLFSAYLVFSMQLGPAMLCAGSVRAKNTMN-IMLTNVLDAAAG--GLFYYLFGFAFAF--  106
Db     4  LISAALVIFMMQPGFALLESGLVRSKNVLNFILMKNFVDLAIGICVLAYVLFGYSLAFGD   63
Qy   107  --GSPSNGFIGKHFF----------GLKD--IPSSSYDYSYFLYQWAFAI--AAAGITS  149
Db    64  SYGEPGNGFIGNGLVWLFLKFLGVSAAGIQDGTLPDG---LPFFLFQLMFAAKTAATIIS  120
Qy   150  GSIAERTQFVAYLIYSSFLTGFVYPVVSHWFWSP--DGWASAFKITDRLFSTGVIDFAGS  207
Db   121  GAVAERIKFSAYLLFSALLGTLVYPPVAHWWGELVGGWLAKLGVLVLILKTKAIDFAGS   180
Qy   208  GVHMVGGIAGLWGALIEGPRMGRF--DHAGRAVALRGHSASLVVLGTFLLWFGWYGFNP   265
Db   181  TVVHIVGGVAGLAAALVLGPRIGRFPDDETGKPEAIRPHNLPFAVLGTFLLWFGWFGFNA  240
Qy   266  GSFNKILLTYGNSGNYYGQWSAVGRTAVTTLAGSTAALTTLFGKRVISGHWNVTDVCNG  325
```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-190-115-100

Query Match        26.8%;   Score 707.5; DB 15;  Length 395;
Best Local Similarity  39.4%;   Pred. No. 3.3e-62;
Matches 172; Conservative  60; Mismatches 147;  Indels  57;  Gaps    12;

Qy    51 YLLFSAYLVFSMQLGFAMLCAGSVRAKNTMNIMLTNVLDAAAGGLFYYLFGFAFAFG-SP 109
         ::||||  ||||| |||| :| ||| :: :|:|||| | |:|| ||| : ||:  |||
Db     2 WILISAALVIFMQPGFALLESGLVRSKNVLNILYKNFQDVAIGVLAYWGFGYSLAFGKSY  61

Qy   110 SNGFIGKHFFGL------KDIPSSSYDYSYFLYQWAFAIAAAGITSGSIAERTQFVAYL 162
         :|||||            | |||: | ||:  |  ||::|| :::|||| :|
Db    62 FSGFIGN--LGLLAAGIQWGTLP---DGLFFLFQLMFAATAITIISGAVAERIKFSAYL 115

Qy   163 IYSSFLTGFVYPVVSHWFWSPDGWASAFKITDRLFSTGVIDFAGSGVVHMVGGIAGLWGA 222
         |:|||| | : || ||: ||  |||  | :|| |||| |||| :||||||||||||
Db   116 LFSALLGTLVYPPVAHWVWGEGGWLAKLGV-------LVDFAGSTVVHIFGGYAGLAAA 167

Qy   223 LIEGPRMGRFDHAGRAVALRGHSASLVVLGTFLLWFGWYGFNPGSFNKILLTYGNSGNYY 282
         |: ||||||| ||: |  |   :  :  || |||||||:||| |||| |  | ||||
Db   168 LVLGPRLGRFT---KNEAITPHNLFPAMLGTLLLWFGWFGFNAGS--------------- 209

Qy   283 GQWSAVGR---TAVTTTLAGSTAALTTLFGKRVISGHWNVTDVCNGLLGGFAAITAGCSV 339
         || ||||    :: || ||  | ||| :  |    | | : | || |||| |||||
Db   210 -ALAADGRARAAAVNTNLAAAGGALTAILISRLKTGKPNMLGLANGALAGLVAITPACGV 268

Qy   340 VEPWAAIVCGFVASIVLIACNK-LAEKVKFDDPLEAAQLHGGCGTWGVIFTALFAKKEYV 398
         :|||  ||: |::| : :  |   |:| ||  |:  |  |||  ||| || :|: :|
Db   269 VSPWGALIIGLIAGVLSVLGYKFLKEKLGIDDPLDVFPVHGVGGIWGGIAVGIFA----- 323

Qy   399 KEVYGLGRAHGLLMGGGGKLLAAHVIQILVIAGW---VSATMGPLFWGLNKLKLLRISSE 455
         |||||| | ||||||||| ||||||||||||||    |: || ||||||| ||  |:
Db   324 -ALYVTSGISGGLLYGNSKQLGVQLIGIAVILAYAFGVTFILGLL---LGLTLGLRVSEE 379

Qy   456 DELAGMDMTRHGGFAY 471
         :|  | ||  :||||
Db   380 EEKVGLDVAEHGETAY 395
```

```
RESULT 8
US-10-369-072-100
; Sequence 100, Application US/10369072
; Publication No. US20040014081A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook II, John P
; APPLICANT: Spaderna, Stephen K
; APPLICANT: Tchernev, Velizar
; APPLICANT: Liu, Xiaohong
; APPLICANT: Shenoy, Suresh
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Taupier, Raymond T
; APPLICANT: Rastelli, Luca
; APPLICANT: Grosse, William M
; APPLICANT: Szerkeres, Edward S
; APPLICANT: Lepley, Denise M
; APPLICANT: Shen, Lei
; APPLICANT: Burgess, Catherine E
; APPLICANT: Shimkets, Richard
; APPLICANT: Padigaru, Muralidhara
; TITLE OF INVENTION: No. US20040014081A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-050 CON2
; CURRENT APPLICATION NUMBER: US/10/369,072
; CURRENT FILING DATE:  2003-02-18
; PRIOR APPLICATION NUMBER: 10/174,372
; PRIOR FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: 09/898,994
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: 60/215,854
; PRIOR FILING DATE: 2000-07-03
```

```
; PRIOR APPLICATION NUMBER: 60/215,856
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/215,902
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/216,585
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/216,722
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/218,622
; PRIOR FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 60/218,992
; PRIOR FILING DATE: 2000-07-17
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 100
; LENGTH: 395
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Domain search
; OTHER INFORMATION: result
US-10-369-072-100

Query Match         26.8%;   Score 707.5; DB 15;  Length 395;
Best Local Similarity  39.4%;   Pred. No. 3.3e-62;
Matches 172; Conservative  60; Mismatches 147;  Indels  57;  Gaps    12;

Qy    51 YLLFSAYLVFSMQLGFAMLCAGSVRAKNTMNIMLTNVLDAAAGGLFYYLFGFAFAFG-SP 109
         ::||||  ||||| |||| :| ||| :: :|:|||| | |:|| ||| : ||:  |||
Db     2 WILISAALVIFMQPGFALLESGLVRSKNVLNILYKNFQDVAIGVLAYWGFGYSLAFGKSY  61

Qy   110 SNGFIGKHFFGL------KDIPSSSYDYSYFLYQWAFAIAAAGITSGSIAERTQFVAYL 162
         :|||||            | |||: | ||:  |  ||::|| :::|||| :|
Db    62 FSGFIGN--LGLLAAGIQWGTLP---DGLFFLFQLMFAATAITIISGAVAERIKFSAYL 115

Qy   163 IYSSFLTGFVYPVVSHWFWSPDGWASAFKITDRLFSTGVIDFAGSGVVHMVGGIAGLWGA 222
         |:|||| | : || ||: ||  |||  | :|| |||| |||| :||||||||||||
Db   116 LFSALLGTLVYPPVAHWVWGEGGWLAKLGV-------LVDFAGSTVVHIFGGYAGLAAA 167

Qy   223 LIEGPRMGRFDHAGRAVALRGHSASLVVLGTFLLWFGWYGFNPGSFNKILLTYGNSGNYY 282
         |: ||||||| ||: |  |   :  :  || |||||||:||| |||| |  | ||||
Db   168 LVLGPRLGRFT---KNEAITPHNLFPAMLGTLLLWFGWFGFNAGS--------------- 209

Qy   283 GQWSAVGR---TAVTTTLAGSTAALTTLFGKRVISGHWNVTDVCNGLLGGFAAITAGCSV 339
         || ||||    :: || ||  | ||| :  |    | | : | || |||| |||||
Db   210 -ALAADGRARAAAVNTNLAAAGGALTAILISRLKTGKPNMLGLANGALAGLVAITPACGV 268

Qy   340 VEPWAAIVCGFVASIVLIACNK-LAEKVKFDDPLEAAQLHGGCGTWGVIFTALFAKKEYV 398
         :|||  ||: |::| : :  |   |:| ||  |:  |  |||  ||| || :|: :|
Db   269 VSPWGALIIGLIAGVLSVLGYKFLKEKLGIDDPLDVFPVHGVGGIWGGIAVGIFA----- 323

Qy   340 VEPWAAIVCGFVASIVLIACNK-LAEKVKFDDPLEAAQLHGGCGTWGVIFTALFAKKEYV 398
         |||||| | ||||||||| ||||||||||||||    |: || ||||||| ||  |:
Db   324 -ALYVTSGISGGLLYGNSKQLGVQLIGIAVILAYAFGVTFILGLL---LGLTLGLRVSBE 379

Qy   456 DELAGMDMTRHGGFAY 471
         :|  | ||  :||||
Db   380 EEKVGLDVAEHGETAY 395
```

```
RESULT 9
US-10-074-978A-250
; Sequence 250, Application US/10074978A
; Publication No. US2004010119A1
; GENERAL INFORMATION:
; APPLICANT: Leite, Mario
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Guo, Xiaojia (Sasha)
; APPLICANT: Fernandes, Elma
; APPLICANT: Li, Li
```

```
                                              Db    59 LP----DGLFFLFQLMFAATATTIISGAVABRIKFSAYLLFSALLGTLVYPPVAHWVWGE 114

; APPLICANT:  Kekuda, Ramesh             Qy   184 DGWASAFKITDRLFSTGVIDFAGSGVVHMVGGIAGLWGALIEGPRMGRFDHAGRAVALRG 243
; APPLICANT:  Liu, Xiahong                             ||: |           :||||||| |||: ||| ||| :  :
; APPLICANT:  Casman, Stacie            Db   115 GGWLAKLGV--------LVDFAGSTVVHIFGGVAGLAAALVLGPRIGRFT---KNEAITP 163
; APPLICANT:  Boldog, Ferenc
; APPLICANT:  Patturajan, Meera         Qy   244 HSASLVVLGTFLLWFGWYGFNPGSFNKILLTYGNSGNYYGQWSAVGR---TAVTTTLAGS 300
; APPLICANT:  Blalock, Angela                          |:  ||||||||| |||||||:: ||
; APPLICANT:  Ballinger, Robert         Db   164 HNLPFAVLGTLLLWFGWFGFNAGS-------------ALTADGRARAAAVNTNLAAA 207
; APPLICANT:  Vernet, Corine
; APPLICANT:  Tchernev, Velizar T       Qy   301 TAALTTLFGKRVISGHWNTVDVCNGLLGGFAAITAGCSVVEPWAAIVCGFVASIVLIACN 360
; APPLICANT:  Maiyankar, Uriel M                       ||: :||  | ||:::| ||:  |:|  :|:  :|:
; APPLICANT:  Gusev, Vladimir           Db   208 GGALTALLISRLKTGKPNMLGLANGALAGLVAITPACGVVSPWGALIIGLIAGVLSVLGY 267
; APPLICANT:  Rastelli, Luca
; APPLICANT:  Mezes, Peter S            Qy   361 KLAEKVKFDDPLEAAQLHGGCGTWGVIFTALFAKKEYVKE--VYGLGRAHGLLMGGGGKL 418
; APPLICANT:  Heyes, Melvin P                          || ||: ||||: :|| ||| ||:          ||| |||
; APPLICANT:  Herrman, John             Db   268 KLKEKLGIDDPLDVFPVHGVGGIWGGIAVGIFAAL-YVNTSGIYG-----GLLY-GNSKQ 320
; APPLICANT:  Pena, Carol E A
; APPLICANT:  Shimkets, Richard A       Qy   419 LAAHVIQILVIAGW---VSATMGPLFWGLNKLKLLRISSEDELAGMDMTRHGGFAY 471
; APPLICANT:  Taupier Jr, Raymond J                    ||:      |||| ||      |||||||||
; APPLICANT:  Moore, No. US2004010119A1Ile  Db   321 LGVQLIGIAVILAYAFGVTFILGLL---LGLTLGLRVSEBEEKVGLDLAEHGETAY 373
; APPLICANT:  Shenoy, Suresh
; APPLICANT:  Edinger, Shlomit          RESULT 10
; APPLICANT:  Gunther, Erik             US-10-033-109-2
; APPLICANT:  Stone, Dave               ; Sequence 2, Application US/10033109
; APPLICANT:  Millet, Isabelle          ; Publication No. US2002014239OA1
; APPLICANT:  Peyman, John              ; GENERAL INFORMATION:
; APPLICANT:  Smithson, Glennda         ;  APPLICANT:  Allen, Stephen M.
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME   ;  APPLICANT:  Rafalski, J. Antoni
; FILE REFERENCE: 21402-269            ;  APPLICANT:  Sakai, Hajime
; CURRENT APPLICATION NUMBER: US/10/074,978A   ;  TITLE OF INVENTION: Nitrogen Transport Metabolism
; CURRENT FILING DATE:  2003-01-07     ;  FILE REFERENCE: BB-1210
; PRIOR APPLICATION NUMBER: 60/268,221 ;  CURRENT APPLICATION NUMBER: US/10/033,109
; PRIOR FILING DATE:  2001-02-12       ;  CURRENT FILING DATE:  2001-12-28
; PRIOR APPLICATION NUMBER: 60/335,109 ;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/384,625
; PRIOR FILING DATE:  2001-10-31       ;  PRIOR FILING DATE:  1999-08-27
; PRIOR APPLICATION NUMBER: 60/312,284 ;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/098,248
; PRIOR FILING DATE:  2001-08-14       ;  PRIOR FILING DATE: EARLIER FILING DATE: 28 August 1998
; PRIOR APPLICATION NUMBER: 60/268,496 ;  NUMBER OF SEQ ID NOS: 14
; PRIOR FILING DATE:  2001-02-13       ;  SOFTWARE: Microsoft Office 97
; PRIOR APPLICATION NUMBER: 60/276,703 ;  SEQ ID NO 2
; PRIOR FILING DATE:  2001-03-16       ;  LENGTH: 183
; PRIOR APPLICATION NUMBER: 60/330,293 ;  TYPE: PRT
; PRIOR FILING DATE:  2001-10-18       ;  ORGANISM: Zea mays
; PRIOR APPLICATION NUMBER: 60/322,127 US-10-033-109-2
; PRIOR FILING DATE:  2001-11-21
; PRIOR APPLICATION NUMBER: 60/280,899  Query Match      23.2%;   Score 613;  DB 13;   Length 183;
; PRIOR FILING DATE:  2001-04-02       Best Local Similarity  65.5%;  Pred. No. 3.4e-53;
; PRIOR APPLICATION NUMBER: 60/310,797 Matches 116;  Conservative  30;  Mismatches 23;  Indels  8;  Gaps  2;
; PRIOR FILING DATE:  2001-08-08
; PRIOR APPLICATION NUMBER: 60/268,646  Qy   330 FAAITAGCSVVEPWAAIVCGFVASIVLIACNKLAEKVKFDDPLEAAQLHGGCGTWGVIFT 389
; PRIOR FILING DATE:  2001-02-14                    ||||||||||||||||:||::|||              :|||||||||||||| |||:|
; Remaining Prior Application data removed - See File Wrapper or PALM.   Db    3 FAAITAGCSVVEPWAAVICGFVSAWVLIGANALAARFRFDDPLEAAQLHGGCGAWGVLFT 62
; NUMBER OF SEQ ID NOS: 547
; SOFTWARE: PatentIn Ver. 2.1          Qy   390 ALFAKKEYVKEVYGLGRAHGLLMGGGGKLLAAHVIQILVIAGWVSATMGPLFWGLNKLKL 449
; SEQ ID NO 250                                    |||::||:||| ::||  |||  ||| |||| |||||| ||||||||||||||||
; LENGTH: 373                          Db   63 GLFARRKYVEEIYGAGRPYGLFMGGGGGKLLAAQIIQILVIAGWVSCTMGPLFYALKLGL 122
; TYPE: PRT
; ORGANISM: Artificial Sequence        Qy   450 LRISSEDELAGMDMTRHGGFAYAYED----DETHKHGMQLR----RVGPNASSTPTT 498
; FEATURE:                                         |||::|:|| | :||   ||||||||||::                ::| :  ||
; OTHER INFORMATION: Description of Artificial Sequence: Ammonium  Db   123 LRISADDEMSGMDLTRHGGFAYVYHDEDPGDKAGVGFMLKSAQNRVEPAAAVAAT 179
; OTHER INFORMATION: Transporter Family domain sequence
US-10-074-978A-250                    RESULT 11
                                      US-10-156-761-13126
Query Match      25.7%;  Score 677.5; DB 15;  Length 373;   ; Sequence 13126, Application US/10156761
Best Local Similarity  40.1%;  Pred. No. 3.1e-59;     ; Publication No. US2003011901BA1
Matches 167; Conservative  53; Mismatches 137; Indels 59;  Gaps 14;   ; GENERAL INFORMATION:
                                      ;  APPLICANT:  OMURA, SATOSHI
Qy   72 GSVRAKNTMNIMLTNVLDAAAGGLFYYLFGPAFAFG-SPSNGFIGKHFFGL------KD 123   ;  APPLICANT:  IKEDA, HARUO
         : ||:|||  :|:     ||           |: ||| ||| ||   :|||| |     ;  APPLICANT:  ISHIKAWA, JUN
Db    1 GLVRSKNVLNLILYKNFQDVAIGVLAYWGFGVSLAFGDSYFSFGIGN--LGLLAAGIQWGT 58   ;  APPLICANT:  HORIKAWA, HIROSHI

Qy   124 IPSSSYDYSYFLYQWAFAIAAAGITSGSIAERTQFVAYLIYSSFLTGFVYPVVSHWFWSP 183
         :| | |:||  ||  |  | |||||||:| ||::| :| ||| ||: |  |||:|||
```

APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATTORI, MASAHIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156,761
CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
PRIOR FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 13126
    LENGTH: 436
    TYPE: PRT
    ORGANISM: Streptomyces avermitilis
US-10-156-761-13126

Query Match   20.3%; Score 535; DB 14; Length 436;
Best Local Similarity   30.8%; Pred. No. 7.6e-45;
Matches 144; Conservative 69; Mismatches 204; Indels 50; Gaps 11;

Qy   41 VDTAFAVDNTYLLFSAYLVFSMQLGFAMLCAGSVRAKNTMNIMLTNVLDAAAGLFYYLF 100
            :|| :||| || :||||:::|:::: : :
Db   8 IDTG---DTAWLLAATALVLLMTPGLALFYGGMVRTKSVLNMLMMSFVSIALVTVVWLAA 64

Qy   101 GFAFAFGSP-SNGFIG--KHF----FGLKDIPSSSYDYSYFLYQWAFAIAAAGITSGSIA 153
            |::| |||:|::|| :   :|||||||||||
Db   65 GYSLAFGDDLAGGLGLGGLDHAGLADIGPADVRGTVPTLFAITFQLTFAIITALISGAIA 124

Qy   154 ERTQFVAYLIYSSFLTGFVIPVVSHWFWSPDGWASAFKITDRLFSTGVIDFAGSGVVHMV 213
            :|:||||| | :||||||||:|||||
Db   125 DRAKFAAWLVFVPLWALLVVVPVAHVWWGPGGW------ILDRL---GALDFAGGLPVEIT 176

Qy   214 GGIAGLWGALIEGPRMGRFDHAGRAVALRGHSASLVVLGTFLLWFGWYGFNPGSFNKILL 273
            ||||||::: :|:||| | :|: |||:|:||||:||||
Db   177 SGASGLALCLVLGPRLG----FKKDAMRPHNLPMVMLGAGLLWFGWFGFNAGS------ 225

Qy   274 TYGNSGNYYGQWSAVGRTAVTTLAGSTAALTTLFGKRVISGHWNVTDVCNGLLGGFAAI 333
            :|::||||||||||| : : :| ::|:||::::|||
Db   226 ALGANG-------LAAAAFLNTLAAGCTGLLGWLFVEQKRDGHPTTLGAASGAVAGLVAI 278

Qy   334 TAGCSVVEPWAAIVCGFVASIVLIACNKLAEKVKFDDPLEAAQLHGGCGTWGVIFTALFA 393
            || |:|| :| : :::| :|:|| :|| : :|
Db   279 TPSCGAVSLLGALVVGVLAAGVVCSYAVGWKFKLNYDDSLDVVGVHLVGGIIGTLLIGVFA 338

Qy   394 KKEYVKEVYGLGRAHGLLMGGGGKLLAAHVIQILVIAGWVSATMGPLFWGLNKLXLLRIS 453
            :| : : :|| : : : ||||||||||||| : ::: ::|:|:: :|
Db   339 AKEMT-----GGTEGILYGGGLGQLGKQLVAVVAVGAYAPAVTYGLGKLLIDKVMGLRAA 392

Qy   454 SEDELAGMDMTRHGGPAYAYEDDETHKGHGMQLRRVGFNASSTPTTDE 500
            :|| :|:|:| |:||
Db   393 -------DHGILGHGAPSSASVVPTAQK 431

RESULT 12
US-10-156-761-10188
; Sequence 10188, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
;   APPLICANT: OMURA, SATOSHI
;   APPLICANT: IKEDA, HARUO
;   APPLICANT: ISHIKAWA, JUN
;   APPLICANT: HORIKAWA, HIROSHI
;   APPLICANT: SHIBA, TADAYOSHI
;   APPLICANT: SAKAKI, YOSHIYUKI
;   APPLICANT: HATTORI, MASAHIRA
;   TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
;   FILE REFERENCE: 249-262
;   CURRENT APPLICATION NUMBER: US/10/156,761
;   CURRENT FILING DATE: 2002-05-29
;   PRIOR APPLICATION NUMBER: JP 2001-204089
;   PRIOR FILING DATE: 2001-05-30

; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 10188
    LENGTH: 446
    TYPE: PRT
    ORGANISM: Streptomyces avermitilis
US-10-156-761-10188

Query Match   18.3%; Score 483; DB 14; Length 446;
Best Local Similarity   28.7%; Pred. No. 1.3e-39;
Matches 129; Conservative 68; Mismatches 187; Indels 66; Gaps 10;

Qy   46 AVDNTYLLFSAYLVFSMQLGFAMLCAGSVRAKNTMNIMLTNVLDAAAGLFYYLFGFAFA 105
            ||| :||| :||||:::|:::: : :|:::|:|
Db   16 AANTGFMLICSALVLIMTPGLAFFYGGMVRKSTLNMLMMSFISMGIVTILWVLYGFSLA 75

Qy   106 FGSPSNGFIG--KHFFGLKDI------PSSSYD-YSYFLYQWAFAIAAAGITSGSIAERT 156
            :||: |: ||:|||
Db   76 FGTDKGSLIGWTSDFVGFTGIGKVELWPGYTIPVYVFATFQLMFAIITPALISGALADRV 135

Qy   157 QFVAYLIYSSFLTGFVYPVVSHWFWSPDGWASAFKITDRLFSTGVIDFAGSGVVHMVGGI 216
            :|:|:::: |:||: || |:||
Db   136 KFTAWSLFVALWATVVYFPVAHVWWGTGGWA--------FDLGVIDFAGGTAVHINAGA 186

Qy   217 AGLWGALIEGPRMGRFDHAGRAVALRGHSASLVVLGTFLLWFGWYGFNPGSFNKILLTYG 276
            :|:|:::: :| ||||:|: |||
Db   187 AALGVILVIGKRVG-----FKKDPMRPHSLPLVMLGCGLLWFGWFGFNAGS------ 232

Qy   277 NSGNYYGQWSAVGRTA-VTTTLAGSTAALTTLFGKRVISGHWNVTDVCNGLLGGFAAITA 335
            :|::||||||||||: :|||:|| ::|:||::::|||
Db   233 ----WLGNDDGVGALMFINTQIATAAAMLAWLAYEKIRHGAFTTLGAASGAVAGLVAITP 288

Qy   336 GCSVVEPWAAIVCGFVASIVLIACNKLAEKVKFDDPLEAAQLHGGCGTWGVIFTALFAKK 395
            ||:|| :| :: :::| :|:| : :: :|
Db   289 SGGAVSPLGAIAVGAIAGVLCAMAVGLKYKFGYDDSLDVVGVHLVGGVAGSLLIGFFASG 348

Qy   396 EYVKEVVGLGRAHGLLMG---------------GGGKLLAAHVIQILVIAGWVSATMGPLFWG 443
            : :|| : : : :||||::|:: ::|
Db   349 G------GQSDATGVFYGDHSFDQLWKQCAGVFAVLAYSLVVSAILAFLIDKTIG------ 397

Qy   444 LNKLKLLRISSEDELAGMDMTRHGGFAYAY 473
            :|:|:::::|||::|:|:|:
Db   398 ------MRVSEDEIAGIDQAEHAETAYDF 421

RESULT 13
US-09-801-368-192
; Sequence 192, Application US/09801368
; Patent No. US20020128250A1
; GENERAL INFORMATION:
;   APPLICANT: Busby, Robert
;   APPLICANT: Cali, Brian
;   APPLICANT: Hecht, Peter
;   APPLICANT: Holtzman, Doug
;   APPLICANT: Madden, Kevin
;   APPLICANT: Maxon, Mary
;   APPLICANT: Milne, Todd
;   APPLICANT: No. US20020128250Alman, Thea
;   APPLICANT: Royer, John
;   APPLICANT: Salama, Sofie
;   APPLICANT: Sherman, Amir
;   APPLICANT: Silva, Jeff
;   APPLICANT: Summers, Eric
;   TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
;   FILE REFERENCE: 109272.147
;   CURRENT APPLICATION NUMBER: US/09/801,368
;   CURRENT FILING DATE: 2001-03-07
;   PRIOR APPLICATION NUMBER: US 09/487,558
;   PRIOR FILING DATE: 2000-01-19
;   PRIOR APPLICATION NUMBER: US 60/160,587
;   PRIOR FILING DATE: 1999-10-20
;   NUMBER OF SEQ ID NOS: 440

```
SOFTWARE: PatentIn version 3.0
SEQ ID NO 192
 LENGTH: 492
 TYPE: PRT
 ORGANISM: Saccharomyces cerevisiae
US-09-801-368-192

Query Match            17.0%;  Score 449.5;  DB 9;  Length 492;
Best Local Similarity  27.9%;  Pred. No. 3.4e-36;
Matches  121; Conservative  77; Mismatches 197; Indels  39; Gaps  10;

Qy     51 YLLFSAYLVFSMQLGFAMLCAGSVRAKNTMNIMLTNVLDAAAGGLFYYLFGFAFAP--GS 108
          ::::||||||:::| |: |||||||:::| |: ||||  :: ||:||
Db     21 FMIIGAALVFFMVPGLGFLIYSGLARRKSALALIWVLMATLVGILQWYFWGYSLAFSKSA 80

Qy    109 PSNGFIGK-HFFGLKDIPSSSYD------YSYFLYQWAFAIAAAGITSGSIAERTQFVAY 161
          ||:| |||  |:||| ||: |::      :||||| :::||| |||  :  ||||||:
Db     81 PNNKFIGNLDSFGFRNVYGKKFDEDAYPELAYATFQMMFSCVNLSIIAGATAERGRLLPH 140

Qy    162 LIYSSFLTGFVYPVVSHWFWSPDGWASAFKITDRLFSTGVIDFAGSGVVHMVGGIAGLWG 221
          ::: |||||||||:|  :||   : ||:||:| :: ||: ||:||| :::
Db    141 MVFLFILATIGYCPVTYWIWSPGGWA-------YQWGVLDWAGGGNIEILSAVSGFVY 191

Qy    222 ALIEGPRMGRFDHAGRAVALRGHSASLVVLGTFLLMFGWYGFNPGSFNKILLTYGNSGNY 281
          ||| ||| |: ||   ::||:| ::: |:| : |:||| |: :|||:
Db    192 SWFLGKRNEKL-----LINFRPHNVSLVTLGTSILNFGWLLLFNSAS-------SLSPNL 238

Qy    282 YGQWSAVGRTAVTTLAGSTAALTTLFGKRVISGHVNVTDVCNGLLGGFAAITAGCSVVE 341
          ||: |:|| : :||||| |  ::  || |||  :||||||||||||||||||:
Db    239 RSVY-AFMNTCLSAITGGMTWCLLDYRSEK----KNSTVGLCSGIISGLVAATPSSGCIT 293

Qy    342 PWAAIVCGFVASIVLIACNKLAEKVKFDDPLEAAQLHGGCGTWGVIFTALFAKKEYVKEV 401
          |  :||||||||:|| ||||:|:| :  | ||:  : ||: ||| |||  || :: :
Db    294 LYGSLIQGIVAGVVCNFATKLKYYAKVDDAMDILAEHGVAGVIGLIFNALFG-ADWVIGM 352

Qy    402 YGLGRAHGLLMGGGGKLLAAHVIQILVIAGWVSATMGPLFWGLNKL--KLLRISSEDELA 459
          ||||| |:||| : | |||||:|| |:|||:||:||||||||:| : ||:||||
Db    353 DGTTEHEGGWVTHNYKQMYKQIAYIAASIGYTAAVTAIICFVLGYIPGMRLRISEEABEA 412

Qy    460 GMDMTRHGGFAYAY 473
          |:|| |:|||||
Db    413 GMDEDQIGFAYDY 426
```

```
 TYPE: PRT
 ORGANISM: Corynebacterium glutamicum
US-09-738-626-6981

Query Match            16.3%;  Score 430.5;  DB 9;  Length 438;
Best Local Similarity  29.2%;  Pred. No. 2.4e-34;
Matches  137; Conservative  68; Mismatches 177; Indels  87; Gaps  17;

Qy     44 AFAVDNTYLLFSAYLVFSMQLGFAMLCAGSVRAKNTMNIMLTNVLDAAAGGLFYLFGFA 103
          | :: ::: ||| || |:: ::|||| |: :|: :|||:::|:|:|:::|
Db      8 AVSGNSAWMLMSASLVLLMTPALALFYGGMSRQKSVLNMWMSFGALGVVTVIYLLWGWS 67

Qy    104 FAFGSPS------NGFIGKHFFGLKD-IPSSSYDY----SY---FLYQWAPAIAAAG 146
          :: ::            :||||||||| ||| :::       :|: ::|: :::|: :
Db     68 MSYGTQSIAGIFANP--EFFGLKDSIVDADGNYIEGAAGYPNIIDIGFQLTFAVISTA 124

Qy    147 ITSGSIAERTQFVAVLIYSSFLTGFVYPVVSHWFW------SPDGWAS-AFKITDRLFS 198
          :|:||||:| :| :|:::::| |||||||||||        :||||:: ||::|||: :
Db    125 LISGALAERVKFSTWLIFSGAWVTLVYFPLAHMVWGGGLLGHNVTGFASWLFGSTDGEAN 184

Qy    199 TGVIDFAGSGVVHMVGGIAGLWGALIEGPRMGRFDHAGRAVALRGHSASLVVLGTFLLWF 258
          ||:::: :: : : |: :|::|| |:: :: :: ||: : |:|::|:: ::::|:||:|
Db    185 IAPIDFAGGTVHISAGTAALVLAFIVGKR---KTFGKAIA-RPHNLPMVMLGAALLWF 239

Qy    259 GWYGFNPGSFNKILLTYGNSGNYYGQWSAVGRTAVTTLAGSTAALTTLFGKRVISGHVN 318
          || ||||| :               :   | || ::|:  ::||:|:|:|||| :|::
Db    240 GWFGFNGGS------AFAADG-----LAGLAWVNTTAATAAAMLGWLATEKFRDGHAT 286

Qy    319 VTDVCNGLLGGFAAITAGCSVVEPWAAIVCGFVASIVLIACNKLAEKVK--FDDPLEAAQ 376
          :|::::|||||||||:||||| :  ||| ::|||||| :|::|||:| :  |||| :::
Db    287 SLGAASGVVAGLVAITPAAGALTPVTSLILGAIGGI-LACLGVGLKYRFGFDDSLDVVG 344

Qy    377 LHGGCGTWGVIFTALFAKKEYVKEVYGLGRAHGLLMGGGGKLLAAHVIQILV------- 428
          : :: | ||: :||| |||:|:|| |||:|: ||||| :: |:||||:|::|
Db    345 VHLVAGLWGTVGVGLLATD-----------IGWFSGGGMDGLKLFIVQIVIALVAVVFA 392

Qy    429 ------IAGWVSATMGPLFWGLNKLKLKLLRISSEDELAGMDMTRGGFAY 471
          :        ::|||:| : : ||  :
Db    393 GVITALIAFALKATVG---W-------RVDDDVEQQGIDTHEHAESAY 430
```

US-09-801-368-194

Query Match   15.7%;   Score 414; DB 9; Length 499;
Best Local Similarity 26.2%; Pred. No. 1.3e-32;
Matches 128; Conservative 73; Mismatches 207; Indels 80; Gaps 17;

Qy   59 VFSMQLGFAMLCAGSVRAKNTMNIMLTNVLDAAAGGLFYLFGFAFAF--GSPSNGFIGK 116
             |: |:| : ||| : : | ::||: ::: :| ::||:|| |||||:||
Db   42 VWIMVPGIGLLYSGLSRKKHALSLLWASMMASAVCIFQWFFWGYSLAFSHNTRGNGFIGT 101

Qy 117 -HFFGLKDI------PSSSYDYSYFLYQWAFAIAAAGTTSGSIAERTQFVAYLIYSSFLT 169
             ||||| ::| | ::: :|||:| || : | :: :::
Db 102 LEFFGFRNVLGAPSSVSSLPDILFAVYQGMFAAVTGALMLGGACERARLFPMMVFLFLWM 161

Qy 170 GFVYPVVSHWFWSPDGWASAFKITDRLFSTGVIDFAGSGVVHMVGGIAGLWGALIEGPRM 229
             | | | |: |:|:||| |: |: || | ||||::
Db 162 TIVYCPIACWVNAEGW-------LVKLGSLDYAGGLCVHLTSGHGGLVYALILGKRN 212

Qy 230 GRFDHAGRAVALRGHSASLVVLGTFLLMFGWYGFNPGSFNKILLTYGNSGN-YYGQWSAV 288
             |: | : | |||||| |||:: || : :||
Db 213 DPVTRKGMP-KYKPHSVTSVVLGTVFLWFGWMFPNGGS--------AGNATIRAWYSI 261

Qy 289 GRTAVTTLAGSTAALTTLF-----GKRVISGHMNVTDVCNGLLGGFAAITAGCSVVEP 342
             :: || | | | ::   :: |:|:::| || ||
Db 262 ----MSTNLAAACGGLTWMVIDYFRCGRK-----WTTVGLCSGIIAGLVGITPAAGFVPI 312

Qy 343 WAAIVCGFVASIVLIACN---KLAEKVKFDDPLEAAQLHGGCGTWGVIFTALFAKKEYVK 399
             |:|:| :: | | || || :|| :|| : :| :|
Db 313 WSAVVIGVVTG--AGCNLAVDLKSLLRIDDGLDCYSIHGVGGCIGSVLTGIFA-ADYVN 368

Qy 400 EVYG--LGRAHGLLMGGGGKLLAAHVIQILVIAGWVSATMGPLFWGLNKLKL--LRISSE 455
             | | | || : || : :|| | : | |:| :| :|: :|: :
Db 369 ATAGSYISPIDGGWINHHYKQVGYQLAGICAALAWTVTVTSILLTWNAIPFLKLRLSAD 428

Qy 456 DELAGMDMTRHGGFAY----AY--------------EDDETHKHGMQLRRVGPN 491
             : | : || :
Db 429 EEELGTDAAQIGEFTYEESTAYIPEPIRSKTSAQMPPPHENIDDKIVGNTDAEKNSTPSD 488

Qy 492 ASSTPTTD 499
             |||| ||
Db 489 ASSTKNTD 496

Search completed: March 9, 2004, 12:01:17
Job time : 41 secs

PM protein - protein search, using sw model

```
                    GenCore version 5.1.6
          Copyright (c) 1993 - 2004  Compugen Ltd.
```

Run on: March 9, 2004, 11:52:48 ; Search time 17 Seconds
                                    (without alignments)
                                    1531.475 Million cell updates/sec

Title: US-10-033-109-4
Perfect score: 2639
Sequence: 1 MSLPACPAEQLAQLLGPNTT.........HGMQLRRVGPNASSTPTTDE 500

Scoring table: BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   SwissProt_42:*

        Pred. No. is the number of results predicted by chance to have a
        score greater than or equal to the score of the result being printed,
        and is derived by analysis of the total score distribution.

                                SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID		Description
1	2129.5	80.7	501	1	AT11_ARATH	P54144	arabidopsis
2	2089.5	79.2	488	1	AT11_LYCES	P58905	lycopersico
3	2011	76.2	519	1	AT13_ARATH	Q9SQH9	arabidopsis
4	1973.5	74.8	504	1	AT14_ARATH	Q9SVT8	arabidopsis
5	1953	74.0	514	1	AT12_LYCES	O04161	lycopersico
6	1919	72.7	514	1	AT12_ARATH	Q9ZPJ8	arabidopsis
7	1591	60.3	460	1	AT13_LYCES	Q9FVN0	lycopersico
8	706	26.8	534	1	AMT1_CAEEL	P54145	caenorhabdi
9	703	26.6	507	1	Y108_SYNY3	P54147	synechocyst
10	678.5	25.7	442	1	YA1_SYNY3	P72935	synechocyst
11	579.5	22.0	687	1	AMT3_CAEEL	Q21565	caenorhabdi
12	569.5	21.6	541	1	Y537_SYNY3	P54148	synechocyst
13	496	18.8	477	1	AMT_MYCTU	Q10968	mycobacteri
14	494	18.7	404	1	NRGA_BACSU	Q07429	bacillus su
15	485	18.4	420	1	YD43_METJA	Q58739	methanococc
16	470	17.8	428	1	AMTB_ECOLI	P37905	escherichia
17	460	17.4	407	1	Y661_METTH	O26757	methanobact
18	457	17.3	423	1	AMT_AQUAE	O66515	aquifex aeo
19	456.5	17.3	412	1	Y663_METTH	O26759	methanobact
20	454.5	17.2	554	1	AMT2_CAEEL	Q20605	caenorhabdi
21	449.5	17.0	492	1	MEP1_YEAST	P40260	saccharomyc
22	446	16.9	489	1	MEP3_YEAST	P53390	saccharomyc
23	436	16.5	391	1	Y058_METJA	Q60366	methanococc
24	414	15.7	499	1	MEP2_YEAST	P41948	saccharomyc
25	366	13.9	475	1	AMT2_ARATH	Q9M6N7	arabidopsis
26	363.5	13.8	452	1	AMT_CORGL	P54146	corynebacte
27	127.5	4.8	455	1	GUDP_BACSU	P42237	bacillus su
28	126	4.8	409	1	RHAG_HUMAN	Q02094	homo sapien
29	124.5	4.7	353	1	RHL_HYLPI	Q28446	hylobates p
30	124	4.7	416	1	RHL_MACMU	Q28849	macaca mula
31	122.5	4.6	444	1	GARP_ECOLI	P42613	escherichia
32	121.5	4.6	456	1	GUDP_PSEPU	P42205	pseudomonas
33	119	4.5	534	1	TNA1_YEAST	P53322	saccharomyc
34	118	4.5	550	1	PUTX_EMENI	P18696	emericella
35	117.5	4.5	416	1	RHLD_GORGO	Q28427	gorilla gor
36	114.5	4.3	499	1	ANSP_ECOLI	P77610	escherichia
37	114	4.3	416	1	RHL_MACFA	Q28481	macaca fac
38	113.5	4.3	416	1	RHLC_GORGO	Q28426	gorilla gor
39	113.5	4.3	497	1	ANSP_SALTY	P40812	salmonella
40	109.5	4.1	439	1	UHPC_ECOLI	P09836	escherichia
41	108	4.1	1094	1	EMBA_MYCTU	P72060	mycobacteri
42	106.5	4.0	416	1	RHD_HUMAN	Q02161	homo sapien
43	106.5	4.0	451	1	YYAJ_BACSU	P37514	bacillus su
44	106.5	4.0	467	1	MMUP_ECOLI	Q47689	escherichia
45	106	4.0	457	1	XYLI_LACBR	O52733	lactobacill

                                ALIGNMENTS

RESULT 1
AT11_ARATH
ID   AT11_ARATH   STANDARD;   PRT;   501 AA.
AC   P54114;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Ammonium transporter 1, member 1 (AtAMT1;1).
GN   AMT1.1 OR AT4G13510 OR T6G15.60.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Landsberg erecta;
RX   MEDLINE=94341258; PubMed=8062823;
RA   Ninnemann O., Janniaux J.-C., Frommer W.B.;
RT   "Identification of a high affinity NH4+ transporter from plants.";
RL   EMBO J. 13:3464-3471(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RX   MEDLINE=20083488; PubMed=10617198;
RA   Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,
RA   Pohl T., Duesterhoeft A., Stiekema W., Entian K.-D., Terryn N.,
RA   Harris B., Ansorge W., Brandt P., Grivell L.A., Rieger M.,
RA   Weichselgartner M., de Simone V., Obermaier B., Mache R., Mueller M.,
RA   Kreis M., Delseny M., Puigdomenech P., Watson M., Schmidtheini T.,
RA   Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,
RA   Vos P., Hoheisel J., Zimmermann W., Wedler H., Ridley P.,
RA   Langham S.-A., McCullagh B., Bilham L., Robben J.,
RA   Van der Schueren J., Grymonprez B., Chuang Y.-J., Vandenbussche F.,
RA   Braeken M., Weltjens I., Voet M., Bastiaens I., Aert R., Defoor E.,
RA   Weitzenegger T., Bothe G., Ramsperger U., Hilbert H., Braun M.,
RA   Holzer E., Brandt A., Peters S., van Staveren M., Dirkse W.,
RA   Mooijman P., Klein Lankhorst R., Rose M., Hauf J., Koetter P.,
RA   Berneiser S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,
RA   De Keyser A., Buysshaert C., Gielen J., Villarroel R., De Clercq R.,
RA   Van Montagu M., Rogers J., Cronin A., Quail M.A., Bray-Allen S.,
RA   Clark L., Doggett J., Hall S., Kay M., Lennard N., McLay K., Mayes R.,
RA   Pettett A., Rajandream M.A., Lyne M., Benes V., Rechmann S.,
RA   Borkova D., Bloecker H., Scharfe M., Grimm M., Loehnert T.-H.,
RA   Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Auer S.,
RA   Gabel C., Fuchs M., Fartmann B., Granderath K., Dauner D., Herzl A.,
RA   Neumann S., Argiriou A., Vitale D., Liguori R., Piravandi E.,
RA   Massenet O., Quigley F., Clabauld G., Muendlein A., Felber R.,
RA   Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,
RA   Chefdor F., Cooke R., Berger C., Monfort A., Casacuberta E.,
RA   Gibbons T., Weber N., Vandenbol M., Bargues M., Terol J., Torres A.,
RA   Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Jesse T.,
RA   Heijnen L., Schwarz S., Scholler P., Heber S., Francs P., Bielke C.,
RA   Frishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,
RA   Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,
RA   Parnell L., Dedhia N., Gnoj L., Schutz K., Huang E., Spiegel L.,

```
Db   119  PTASADYSNFLYQWAFAIAAAGITSGSIAERTQFVAYLIYSSFLTGFVPVVSHWFWSVD 178
                ::|||  |||||||||||||||||||||||||||||||||||||||||||||
Qy   185  GWASAFKIT-DRLFSTGVIDFAGSGVVHMVGGIAGLWGALIEGPRMGRFDHAGRAVALRG 243
          |||| :  ||||||||:|||||||||||||||||||||||||||:|||:|| |||:||||
Db   179  GWASPRTDGDLLFSTGAIDFAGSGVVHMVGGIAGLWGALIEGPRLGRFDNGGRAIALRG 238
Qy   244  HSASLVLGTFLLWFGWYGFNPGSFNKILLTYGNSGNYYGQWSAVGRTAVTTLLAGSTAA 303
          |||||||||||||||||||||||||||||:||  ||:|||||||||||||||||||:|||
Db   239  HSASLVLGTFLLWFGWYGFNPGSFNKILVTY-ETGTYNGQWSAVGRTAVTTLLAGCTAA 297
Qy   304  LTTLFGKRVIGGHWNVTDVCNGLLGGFAAITAGCSVVBPWAAIVCGFVASIVLIACNKLA 363
          |||||||||: |||||||||||||||||||| |||||||||||:|||||  |  ||||||
Db   298  LTTLFGKRLLSGHWNVTDVCNGLLGGFAAITGGCSVVBPWAAIICGFVAALVLLGCNKLA 357
Qy   364  EKVKFDDPLEAAQLHGGCGTWGVIFTALFAKKEVKEVVG--LGRAHGLLMGGGGKLLAA 421
          || :|||||||||||||||| ||:|||||| :  :  |   ||| ||| ||||||||||
Db   358  EKLKYDDPLEAAQLHGGCGAWGLIFTALFAQERYLNQIYGNKFGRPHGLFMGGGGKLLGA 417
Qy   422  HVIQILVIAGWVSATMGPLFWGLNKLKLLRISSEDELAGMDMTRHGGFAYAY-EDDETHK 480
           :|||:| ||||||||| |:  :|:|||||||||||:|||||||||||||:|| :||:||
Db   418  QLIQIIVITGWVSATMGTLFFIIKKMKLLRISSEDEMAGMDMTRHGGFAYMYFDDDESHK 477
Qy   481  HGMQLRRVGP-----NASSTPT 497
            :||||  ||     :||:|||
Db   478  -AIQLRVEPRSPSPSGANTTPT 499
```

```
RESULT 2
AT11_LYCES

ID   AT11_LYCES     STANDARD;     PRT;     488 AA.
AC   P58905;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Ammonium transporter 1, member 1 (LeAMT1;1).
GN   AMT1.1.
OS   Lycopersicon esculentum (Tomato).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC   lamiids; Solanales; Solanaceae; Solanum.
OX   NCBI_TaxID=4081;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96353959; PubMed=8755617;
RA   Lauter F.-R., Ninnemann O., Bucher M., Riesmeier J.W., Frommer W.B.;
RT   "Preferential expression of an ammonium transporter and of two
RT   putative nitrate transporters in root hairs of tomato.";
RL   Proc. Natl. Acad. Sci. U.S.A. 93:8139-8144(1996).
RN   [2]
RP   INDUCTION.
RX   MEDLINE=20208060; PubMed=10743657;
RA   von Wiren N., Lauter F.-R., Ninnemann O., Gillissen B., Walch-Liu P.,
RA   Engels C., Jost W., Frommer W.B.;
RT   "Differential regulation of three functional ammonium transporter
RT   genes by nitrogen in root hairs and by light in leaves of tomato.";
RL   Plant J. 21:167-175(2000).
CC   -!- FUNCTION: Involved in high affinity ammonium uptake.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC   -!- TISSUE SPECIFICITY: Root hairs and leaves.
CC   -!- INDUCTION: By nitrogen deficiency.
CC   -!- SIMILARITY: Belongs to the ammonium transporter (TC 2.A.49)
CC       family.
```

```
DR   EMBL; X92854; -; NOT_ANNOTATED_CDS.
```

```
RA   Sekhon M., Murray J., Sheet P., Cordes M., Abu-Threideh J.,
RA   Stoneking T., Kalicki J., Graves T., Harmon G., Edwards J.,
RA   Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
RA   Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,
RA   Kramer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L.W.,
RA   Nelson J., Spieth J., Ryan E., Andrews S., Geisel C., Layman D.,
RA   Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshu C.,
RA   Antoniou B., Zidanic M., Strong C., Sun H., Lamar B., Yordan C.,
RA   Ma P., Zhong J., Preston R., Vil D., Shekher M., Matero A., Shah R.,
RA   Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Till S.,
RA   Granat S., Shohdy N., Hasegawa A., Hameed A., Lodhi M., Johnson A.,
RA   Chen E., Marra M.A., Martienssen R., McCombie W.R.;
RT   "Sequence and analysis of chromosome 4 of the plant Arabidopsis
RT   thaliana.";
RL   Nature 402:769-777(1999).
RN   [3]
RP   TISSUE SPECIFICITY.
RC   STRAIN=cv. C24;
RX   MEDLINE=9264313; PubMed=10330477;
RA   Gazzarini S., Lejay L., Gojon A., Ninnemann O., Frommer W.B.,
RA   von Wiren N.;
RT   "Three functional transporters for constitutive, diurnally regulated,
RT   and starvation-induced uptake of ammonium into Arabidopsis roots.";
RL   Plant Cell 11:937-948(1999).
CC   -!- FUNCTION: Involved in high affinity ammonium uptake.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC   -!- TISSUE SPECIFICITY: High expression in root.
CC   -!- SIMILARITY: Belongs to the ammonium transporter (TC 2.A.49)
CC       family.
```

```
DR   EMBL; X75879; CAA53473.1; -.
DR   EMBL; AL049656; CAB41109.1; -.
DR   EMBL; AL161536; CAB78393.1; -.
DR   PIR; T06653; T06653.
DR   InterPro; IPR001905; Ammonium_transpt.
DR   Pfam; PF00909; Ammonium_transp; 1.
DR   TIGRFAMs; TIGR00836; amtB; 1.
DR   PROSITE; PS01219; AMMONIUM_TRANSP; 1.
KW   Transport; Transmembrane; Multigene family.
FT   TRANSMEM      8     28       POTENTIAL.
FT   TRANSMEM     46     66       POTENTIAL.
FT   TRANSMEM     81    101       POTENTIAL.
FT   TRANSMEM    128    148       POTENTIAL.
FT   TRANSMEM    152    172       POTENTIAL.
FT   TRANSMEM    199    219       POTENTIAL.
FT   TRANSMEM    243    263       POTENTIAL.
FT   TRANSMEM    311    331       POTENTIAL.
FT   TRANSMEM    333    353       POTENTIAL.
FT   TRANSMEM    366    386       POTENTIAL.
FT   TRANSMEM    420    440       POTENTIAL.
SQ   SEQUENCE   501 AA;  53576 MW;  070BAF8228302BFF CRC64;

Query Match   80.7%;  Score 2129.5;  DB 1;  Length 501;
Best Local Similarity  80.9%;  Pred. No. 1.6e-141;
Matches 407;  Conservative 42;  Mismatches 39;  Indels 15;  Gaps 7;

Qy    5  ACPAEQLAQLLGPNTTDASAAASLICGHFAAVDSKFVDTAFAVDNTYLLFSAYLVFSMQL 64
            :| :|||||||||  :| :|| |:|||::|||||||||||||||||||||||
Db    2  SCSATDLAVLLGPN--ATAAANYICGQLGDVNNKFIDTAFAIDNTYLLFSAYLVFSMQL 58
Qy   65  GFAMLCAGSVRAKNTMNIMLTNVLDAAAGLFYLLFGFAFAFGSPSNGFIGKHFPGLKDI 124
         |||||||||||||||||||||||||||||||||:||||||||||||:||:||||||
Db   59  GFAMLCAGSVRAKNTMNIMLTNVLDAAAGLFYILFGYAFAFGSPSNGFIGKHYFGLKDI 118
Qy  125  PSSSYDYSYFLYQWAFAIAAAGITSGSIAERTQFVAYLIYSSFLTGFVPVVSHWFWSPD 184
```

DR   InterPro; IPR001905; Ammonium_transpt.
DR   Pfam; PF00909; Ammonium_transp; 1.
DR   TIGRFAMs; TIGR00836; amt; 1.
DR   PROSITE; PS01219; AMMONIUM_TRANSP; 1.
KW   Transport; Transmembrane; Multigene family.
FT   TRANSMEM   47   69   POTENTIAL.
FT   TRANSMEM   90   109   POTENTIAL.
FT   TRANSMEM   129   148   POTENTIAL.
FT   TRANSMEM   153   175   POTENTIAL.
FT   TRANSMEM   195   217   POTENTIAL.
FT   TRANSMEM   238   257   POTENTIAL.
FT   TRANSMEM   281   303   POTENTIAL.
FT   TRANSMEM   316   333   POTENTIAL.
FT   TRANSMEM   337   356   POTENTIAL.
FT   TRANSMEM   368   387   POTENTIAL.
FT   TRANSMEM   418   440   POTENTIAL.
SQ   SEQUENCE   488 AA; 52356 MW; C41E48445B0D2FE8 CRC64;

Query Match   79.2%; Score 2089.5; DB 1; Length 488;
Best Local Similarity   80.5%; Pred. No. 9.4e-139;
Matches 396; Conservative 39; Mismatches 50; Indels 7; Gaps 4;

Qy   5 ACPAEQLAQLLGPNTTDASAAASLICGHFAAVDSKFVDTAFAVDNTYLLFSAYLVFSMQL 64
           |||||:|| |||||||| :: ||||:|||||||: :|:|||||||||||||||||||||
Db   2 ACSVDTLAPFLGPNTTNVAAAASYICNQPSGVSDRFVDTGYAIDSTYLLFSAYLVFSMQL 61

Qy   65 GFAMLCAGSVRAKNTMNIMLTNVLDAAAGGLFYLFGFAFAFGSPSNGFIGKHFFGLKDI 124
           ||||||||||||||||||||||||||||||||||||||||:||| |||||||||||||
Db   62 GFAMLLAGSVR--NTMNIMLTNVLDAAAGGLFYYLFGFAFALGGPSNGFIGRHFFGLKEI 119

Qy   125 PSSSYDYSYFLYQWAFAIAAAGITSGSIAERTQFVAYLIYSSFLTGFVYPVVSHWFWSPD 184
           |:||||||||||||||||||||||||||||||||||||||||||||||||||||||||:
Db   120 PSNSFDYMNFLYQWAFAIAAAGITSGSIAERTQFVAYLIYSSFLTGFVYPVVSHWFWTPD 179

Qy   185 GWASAFKITDRLFSTGVIDFAGSGVVHMVGGIAGLWGALIEGPRMGRFDHAGRAVALRGH 244
           |||| |: :||||||||||||||||||||||||||||||:|||||||:|:|||||||||
Db   180 GWASPTN-SNLLFGSGVIDFAGSGVVHMVGGIAGFYGALIEGPRIGRYDHTGRSVALRGH 238

Qy   245 SASLVVLGTFLIWFGWYGFNPGSFNKILLTYGNSGNYYGQWSAVGRTAVTTLAGSTAAL 304
           ||||||||||||||||||||||||:|||||||:||:|||||||||||||||||||||||
Db   239 SASLVVLGTFLLWFGWYGFNPGSFNKILVLVTYGASGCYYGQWSAVGRTAVTTTLAGCTAAL 298

Qy   305 TTLFGKRVISGHWNVTDVCNGLLGGFAATTAGCSVVEPWAAIVCGFVASIVLIACNKLAE 364
           ||||||:|||||||||||||||||||||||||:|||||||||||||:: ||||:||||
Db   299 TTLFGKRILSGHWNVTDVCNGLLGGFAAITAGCSVVEPWAAIICGFVAALVLIGFNMLAE 358

Qy   365 KVKFDDDPLEAAQLHGGCGTWGVIFTALFAKKEYVKEVY--GLGRAHGLLMGGGGKLLAAH 422
           |:|||||||||||||||| |:|||| || ::| || :||||||||||||||:||||
Db   359 KFKYDDDPLEAAQLHGGCGAWGIIFTGLFAKGEFVDQVYPGKPGRPHGLFMGGGGKLLGAH 418

Qy   423 VIQILVIAGWSATMGPLFWGLNKLKLLRISSEDELAGMDMTRHGGFAYAYEDDETHKHG 482
           ||||:||||||||||||||||:||||||||||||:||||:||||||| |||||: :||
Db   419 IIQILVIIGWVSATMGPLFYILHKFKLLRISSEDEMAGMDLTRHGGFAYYHEEDP--KLG 476

Qy   483 MQLRRVGPNASS 494
           ||::|: |:
Db   477 MQMRRIEPTTST 488

RESULT 3
AT13_ARATH
ID   AT13_ARATH   STANDARD;   PRT;   519 AA.
AC   Q9SQH9; Q9SLK15;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Ammonium transporter 1, member 3 (AtAMT1;3).
GN   AMT1.3 OR AT3G24300 OR K7M2.6.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;

RN   [1]
RP   SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RC   STRAIN=cv. C24; TISSUE=Seed;
RX   MEDLINE=99264313; PubMed=10330477;
RA   Gazzarrini S., Lejay L., Gojon A., Ninnemann O., Frommer W.B.,
RA   von Wiren N.;
RT   "Three functional transporters for constitutive, diurnally regulated,
RT   and starvation-induced uptake of ammonium into Arabidopsis roots.";
RL   Plant Cell 11:937-948(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RX   MEDLINE=20363099; PubMed=10907853;
RA   Kaneko T., Katoh T., Sato S., Nakamura A., Asamizu E., Tabata S.;
RT   "Structural analysis of Arabidopsis thaliana chromosome 3. II.
RT   Sequence features of the 4,251,695 bp regions covered by 90 P1, TAC
RT   and BAC clones.";
RL   DNA Res. 7:217-221(2000).
CC   -!- FUNCTION: Involved in high affinity ammonium uptake.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC   -!- TISSUE SPECIFICITY: High expression in root.
CC   -!- INDUCTION: Highest expression at the end of the light period.
CC   -!- SIMILARITY: Belongs to the ammonium transporter (TC 2.A.49)
CC   family.
CC   --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --
DR   EMBL; AF083035; AAD54638.1; -.
DR   EMBL; AP000382; BAB02929.1; -.
DR   InterPro; IPR001905; Ammonium_transpt.
DR   Pfam; PF00909; Ammonium_transp; 1.
DR   TIGRFAMs; TIGR00836; amt; 1.
DR   PROSITE; PS01219; AMMONIUM_TRANSP; 1.
KW   Transport; Transmembrane; Multigene family.
FT   TRANSMEM   50   70   POTENTIAL.
FT   TRANSMEM   85   105   POTENTIAL.
FT   TRANSMEM   131   151   POTENTIAL.
FT   TRANSMEM   156   176   POTENTIAL.
FT   TRANSMEM   203   223   POTENTIAL.
FT   TRANSMEM   247   267   POTENTIAL.
FT   TRANSMEM   315   335   POTENTIAL.
FT   TRANSMEM   337   357   POTENTIAL.
FT   TRANSMEM   370   390   POTENTIAL.
FT   TRANSMEM   423   443   POTENTIAL.
FT   TRANSMEM   493   513   POTENTIAL.
FT   CONFLICT   494   519   LLLAFNFQLFGNLLPFKYCLGFGFEI -> TPPRV (IN
FT   REF. 2).
SQ   SEQUENCE   519 AA; 55755 MW; D9E3A66A3EF70A6E CRC64;

Query Match   76.2%; Score 2011; DB 1; Length 519;
Best Local Similarity   77.8%; Pred. No. 3e-133;
Matches 381; Conservative 38; Mismatches 59; Indels 12; Gaps 5;

Qy   6 CPAEQLAQLLGPNTTDASAAASLICGHFAAVDSKFVDTAFAVDNTYLLFSAYLVFSMQLG 65
           ||||||||||||| :||||||||||||||||:| ::|||||||||||||||||||||||
Db   7 CSAADLATLLGPN--ATRAADYICGQLGTVNNKFTDAAFAIDNTYLLFSAYLVFAMQLG 63

Qy   66 FAMLCAGSVRAKNTMNIMLTNVLDAAAGGLFYLFGFAFAFGSPSNGFTGKHFFGLKDIP 125
           ||||||||||| |||||||||||||||||||||||||||||:: :|||||||||||||
Db   64 FAMLCAGSVRAKNTMNIMLTNVLDAAAGGLFYLFGYAFAFGSSEGFTGRHNFALRDFP 123

Qy   126 SSSYDYSYFLYQWAFAIAAAGITSGSIAERTQFVAYLIYSSFLTGFVYPVVSHWFWSPDG 185
           ::: ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   124 TPTADYSFFLYQWAFAIAAAGITSGSIAERTQFVAYLIYSSFLTGFVYPVVSHWFWSPDG 183

Qy   186 WASAFK-ITDRLFSTGVIDFAGSGVVHMVGGIAGLWGALIEGPRMGRFDHAGRAVALRGH 244
           |||:|: |||||||||||||||||||||||||||||||||||||||||| |||:||||

```
Db  184 WASPFRSADDRLFSTGAIDFAGSGVHMVGIAGLWGALIEGPRRGRFEKGGRAIALRGH 243
        |    | ||    ||  ||||||||| ||| |||||||||||||  || | |||   ||
Qy  245 SASLVVLGTFLLWFGWYGFNPGSFNKILLTYGNSGNYYGQWSAVGRTAVTTLAGSTAAL 304
        |||||||||||||||||||||||| |||| ||  |||| ||| |||||| | |  ||||
Db  244 SASLVVLGTFLLWFGWYGFNPGSFTKILLVPY-NSGSNYGQWSGIGRTAVNTTLSGCTAAL 302

Qy  305 TTLFGKRVISGHWNVTDVCNGLLGGFAAITAGCSVVEPWAAIVCGFVASIVLIACNKLAE 364
        ||||||| :||||||||||||||||||||||||||||||||||||||:||||| |||||
Db  303 TTLFGKRLLSGHWNVTDVCNGLLGGFAAITAGCSVVEPWAAIVCGFMASVVLIGCNKLAE 362

Qy  365 KVKFDDPLEAAQLHGGCGTWGVIFTALFAKKEYVKEVYGL--GRAHGLLMGGGGKLLAAH 422
        :| :|||||||||||||| ||::| ||||| :|::||||    || ||:||||||||:|
Db  363 LVQYDDPLEAAQLHGGCGAWGLLFVGLFAKEKYLNEVYGATPGRPYGLFMGGGGKLLGAQ 422

Qy  423 VIQILVIAGWVSATMGPLFWGLNKLKLLRISSEDELAGMDMTRHGGFAYAYEDDETHKHG 482
        ::||||| ||||||||.|| | |::||||||:::  |||||||||||||:| ||:| ||
Db  423 LVQILVIVGWVSATMGTLFFILKRINLLRISEQHEMQGMDMTRHGGFAYIYHDNDDESH- 481

Qy  483 MQLRRVGPNA 492
        ||   |
Db  482 ----RVDPGS 487

RESULT 4
AT14_ARATH
ID   AT14_ARATH    STANDARD;    PRT;    504 AA.
AC   Q9SVT8;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Ammonium transporter 1, member 4 (AtAMT1;4).
GN   AMT1.4 OR AT4G28700 OR F16A16.190.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   euroids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RX   MEDLINE=20083488; PubMed=10617198;
RA   Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,
RA   Pohl T., Duesterhoeft A., Stiekema W., Entian K.-D., Terryn N.,
RA   Harris B., Ansorge W., Brandt P., Grivell L.A., Rieger M.,
RA   Weichselgartner M., de Simone V., Obermaier B., Mache R., Mueller M.,
RA   Kreis M., Delseny M., Puigdomenech P., Watson M., Schmidtheini T.,
RA   Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,
RA   Vos P., Hoheisel J., Zimmermann W., Wedler H., Ridley P.,
RA   Langham S.-A., McCullagh B., Bilham L., Robben J.,
RA   Van der Schueren J., Grymonprez B., Chuang Y.-J., Vandenbussche F.,
RA   Braeken M., Weltjens I., Voet M., Bastiaens I., Aert R., Defoor E.,
RA   Weitzenegger T., Bothe G., Ramsperger U., Hilbert H., Braun M.,
RA   Holzer E., Brandt A., Peters S., van Staveren M., Dirkse W.,
RA   Mooijman P., Klein Lankhorst R., Rose M., Hauf J., Koetter P.,
RA   Berneiser S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,
RA   De Keyser A., Buysshaert C., Gielen J., Villarroel R., De Clercq R.,
RA   Van Montagu M., Rogers J., Cronin A., Quail M.A., Bray-Allen S.,
RA   Clark L., Doggett J., Hall S., Kay M., Lennard N., McLay K., Mayes R.,
RA   Pettett A., Rajandream M.A., Lyne M., Benes V., Rechmann S.,
RA   Borkova D., Bloecker H., Scharfe M., Grimm M., Loehnert T.-H.,
RA   Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Auer S.,
RA   Gabel C., Fuchs M., Fartmann B., Granderath K., Dauner D., Herzl A.,
RA   Neumann S., Argiriou A., Vitale D., Liguori R., Piravandi E.,
RA   Massenet O., Quigley F., Clabauld G., Muendlein A., Felber R.,
RA   Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,
RA   Chefdor F., Cooke R., Berger C., Monfort A., Casacuberta E.,
RA   Gibbons T., Weber N., Vandenbol M., Bargues M., Terol J., Torres A.,
RA   Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Jesse T.,
RA   Heijnen L., Schwarz S., Scholler P., Heber S., Francs P., Bielke C.,
RA   Frishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,
RA   Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,
RA   Parnell L., Dedhia N., Gnoj L., Schutz K., Huang E., Spiegel L.,
RA   Sekhon M., Murray J., Sheet P., Cordes M., Abu-Threiden J.,
RA   Stoneking T., Kalicki J., Graves T., Harmon G., Edwards J.,
RA   Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
RA   Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,
RA   Kramer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L.W.,
RA   Nelson J., Spieth J., Ryan E., Andrews S., Geisel C., Layman D.,
RA   Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshu C.,
RA   Antonoiu B., Zidanic M., Strong C., Sun H., Lamar B., Yordan C.,
RA   Ma P., Zhong J., Preston R., Vil D., Shekher M., Matero A., Shah R.,
RA   Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman M., Till B.,
RA   Granat S., Shohdy N., Hasegawa A., Hameed A., Lodhi M., Johnson A.,
RA   Chen E., Marra M.A., Martienssen R., McCombie W.R.;
RT   "Sequence and analysis of chromosome 4 of the plant Arabidopsis
RT   thaliana.";
RL   Nature 402:769-777(1999).
CC   -!- FUNCTION: Involved in high affinity ammonium uptake.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC   -!- SIMILARITY: Belongs to the ammonium transporter (TC 2.A.49)
CC       family.
CC   --------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------
DR   EMBL; AL035355; CAA22982.1; -.
DR   EMBL; AL161573; CAB81458.1; -.
DR   PIR; T04529; T04529.
DR   InterPro; IPR001905; Ammonium_transport.
DR   Pfam; PF00909; Ammonium_transp; 1.
DR   TIGRFAMs; TIGR00836; amt; 1.
DR   PROSITE; PS01219; AMMONIUM_TRANSP; 1.
KW   Transport; Transmembrane; Multigene family.
FT   TRANSMEM     12     32      POTENTIAL.
FT   TRANSMEM     55     75      POTENTIAL.
FT   TRANSMEM     90    110      POTENTIAL.
FT   TRANSMEM    136    156      POTENTIAL.
FT   TRANSMEM    161    181      POTENTIAL.
FT   TRANSMEM    207    227      POTENTIAL.
FT   TRANSMEM    251    271      POTENTIAL.
FT   TRANSMEM    318    338      POTENTIAL.
FT   TRANSMEM    344    364      POTENTIAL.
FT   TRANSMEM    377    397      POTENTIAL.
FT   TRANSMEM    430    450      POTENTIAL.
FT   DOMAIN       20     29      POLY-ALA.
SQ   SEQUENCE   504 AA;   53657 MW;   089B12A9EE06FCDA CRC64;

Query Match     74.8%;   Score 1973.5;   DB 1;   Length 504;
Best Local Similarity   75.3%;   Pred. No. 1.2e-130;
  Matches 374;   Conservative 45;   Mismatches 67;   Indels 11;   Gaps 5;

Qy    5 ACPAEQLAQLL--GPNTTDASAAASLICGHFAAVDSKFVDTAFAVDNTYLLFSAYLVFSM 62
         : |  | ||  ||||| || |:|| ||| :|:||  || ||:|:||:|||||||||| :|
Db    6 SCSASDLIPLLSGGANATARAAAAEYICGRFDTVAGKFTDAAYAIDNTYLLFSAYLVFAM 65

Qy   63 QLGFAMLCAGSVTRAKNTMNIMLTNVLDAAAGGLFYYLFGFAFGSPSNGFIGKHPFGLK 122
         ||||||||||||||||||||||||||:|||||||||||||||||||||||||||||| :
Db   66 QLGFAMLCAGSVTRAKNTMNIMLTNVIDAAAGGLFYYLFGFAFGSPSNGFIGKHPFGMY 125

Qy  123 DIPSSSYDYSYFLYQWAFAIAAGITSGSIAERTQFVAVLIYSSFLTGFVYPVVSHWFWS 182
         | :  :: ||||||| | || ||||||||||||||| ::|||||||| ||:|||||||
Db  126 DFPQPTFDYPYFLYQWFAIAAAGITSGSIAERTQFVAVLIYSSFLTGLVYPIVSHWFWS 185

Qy  183 PDGWASAFKITDRLFSTGVIDFAGSGVVHMVGGIAGLWGALIEGPRMGRPFDHAGRAVALR 242
         .||||| :  :  ||:||||||||||||||||||||||||||||||:||||| | | |||
Db  186 SDGWASPARSENLLFQSGVIDFAGSGVVHMVGGIAGLWGALIEGPRIGRFGVGGKPVTLR 245

Qy  243 GHSASLVVLGTFLLWFGWYGFNPGSFNKILLTYGNS--GNYYGQWSAVGRTAVTTTLAGS 300
         ||||:|||||||||||||||||||||  |:|  || :  :  ||||||||||||||||||
Db  246 GHSATLVVLGTFLLWFGWYGFNPGSFATIFKAYGETPGSSFYGQWSAVGRTAVTTTLAGC 305
```

```
FT   TRANSMEM       428    448       POTENTIAL.
SQ   SEQUENCE       514 AA;  55543 MW;  9987FB9FE17173C9 CRC64;

     Query Match             74.0%;  Score 1953;  DB 1;  Length 514;
     Best Local Similarity   75.3%;  Pred. No. 3.3e-129;
     Matches   379;  Conservative   43;  Mismatches   71;  Indels   10;  Gaps    7;

Qy          2  SLPACPAEQLAQLLGPNTTDASAAASLICGHFAAVDSKFVDTAFAVDNTYLLFSAYLVFS   61
               |  :: ::| ::|  ::|:.:
Db          3  SAMTCSAAELFPHLG-SSANATAAEFICSRFSAVSEYLTNTTYAVDTTYLLFSAYIVFA   61

Qy         62  MQLGFAMLCAGSVRAKNTMNIMLTNVLDAAAGLFYLFGFAFAFGSPSNGFIGKHFFGL   121
               |||||||||||||||||||:|::||||||||||||||||||: |||| |:||||||||
Db         62  MQLGFAMLCAGSVRAKNTMNIMLTNVLDAAAGGFSYYLFGFAFAFGAPSNGFIGKHFFGL   121

Qy        122  KDIPSSSYDYSYFLYQWAFAIAAAGITSGSIAERTQFVAVLIYSSFLTGFVYPVVSHWFW   181
               : | ||:||:|||||||||||||||||||||||||||||||||||||||||||||||||
Db        122  KEFPSQAFDYSYFLYQWAFAIAAAGITSGSIAERTQFVAVLIYSSFLTGFVYPIVSHWFW   181

Qy        182  SPDGWASAFKITD--RLFSTGVIDFAGSGVVHMVGGIAGLWGALIEGPRMGRFDHAGRAV   239
               |||||| |  |    |:||||||||||||||||||||||||| :| : :|:
Db        182  SGDGWASASK-TDGNLLLRFGVIDFAGSGVVHMVGGIAGLWGAFIEGPRIGRFDRSGRSV   240

Qy        240  ALRGHSASLVVLGTFLLWFGWYGFNPGSFNKILLTYGNS--GNYYGQWSAVGRTAVTTTL   297
               |||||||||||||||||||||||||||||||||||||||  ||||||||| | |||||||
Db        241  ALRGHSASLVVLGTFLLWFGWYGFNPGSFLTILKSYDHTIRGTYYGQWSAIGRTAVTTTL   300

Qy        298  AGSTAALTTLFGKRVISGHWNVTDVCNGLLGGFAITAGCSVVEPWAAIVCGFVASIVLI   357
               ||||||||||||: ||
Db        301  AGTTAALTTLFCKRLLVAHWNVDVCNGLLGGFAITSGCAVVEPWAAIVCGFIAAWVLI   360

Qy        358  ACNKLAEKVKFDDPLEAAQLHGGCGTWGVIFTALFAKKBYVKEVY-GL-GRAHGLLMGGG   415
               | : |:|: || | :|||:||||||| |:|:|||: |||||||||  ||  |||||||
Db        361  GFNALAAKLKYDDPLEAAQLHGGCGSWGIIFTGLFAKKEYNEVYPGFPNRPYGLFMGGG   420

Qy        416  GKLLAAHVIQILVIAGWVSATMGPLFWGLNKLKLLRISSEDELAGMDMTRHGGFAYAY---473
               :||| |::||| |||||||||||||||||| |||| || |||: |||||||||||||
Db        421  GKLLGAQVIQVVVIIGWVSVTMGPLFYLLHKFKLLRISRDDETAGMDLTRHGGFAYIYHD  480

Qy        474  EDDETHKHGMQLRRVGPNASSTP  496
               || |:|| ||:||||:| |||
Db        481  EDEGSSMPGFKMTRVEFTNTSTP  503


RESULT 5
AT12_LYCES

ID   AT12_LYCES     STANDARD;      PRT;   514 AA.
AC   Q9ZPJ8; Q9SQH8;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Ammonium transporter 1, member 2 (AtAMT1;2).
GN   AMT1.2 OR AT1G64780 OR F13011.9.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RC   STRAIN=cv. C24; TISSUE=Seed;
RX   MEDLINE=9926313; PubMed=10330477;
RA   Gazzarrini S., Lejay L., Gojon A., Ninnemann O., Frommer W.B.,
RA   von Wiren N.;
RT   "Three functional transporters for constitutive, diurnally regulated,
RT   and starvation-induced uptake of ammonium into Arabidopsis roots.";
RL   Plant Cell 11:937-948(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. C24; TISSUE=Root;
RA   Shelden M.C., Howitt S.M., Udvardi M.K.;
RT   "Arabidopsis thaliana AtAMT1;2 from N-deprived roots.";
RL   Submitted (DEC-1998) to the EMBL/GenBank/DDBJ databases.
```

```
Qy        301  TAALTTLFGKRVISGHWNVTDVCNGLLGGFAITAGCSVVEPWAAIVCGFVASIVLIACN  360
               ||||||||||||||:||||||||||||||||||||||||||||||||||||:||: ||
Db        306  TAALTTLFGKRLIDGYWNVTDVCNGLLGGFAITSGCSVVEPWAAIVCGFVRAWVLMGCN  365

Qy        361  KLAEKVKFDDPLEAAQLHGGCGTWGVIFTALFAKKEYVKEVYG--LGRAHGLLMGGGGKL  418
               ::||::|||||||||||||||||||  ||:| |  |: :|:  |   ||||||||||||:|
Db        366  RLAEKLQFDDPLEAAQLHGGCGAWGIIFTGLFAEKRYIAEIFGGDPNRPFGLLMGGGGRL  425

Qy        419  LAAHVIQILVIAGWVSATMGPLFWGLNKLKLKLLRISSEDELAGMDMTRHGGFAYAYEDDET  478
               |||||||||||||||||  ||  | |||||:|:|||||||||||||||||   |: |
Db        426  LAAHVVQILVITGWVSVTMGTLFFILHKLKLLRIPAEDEIAGVDPTSHGGLAYMYTEDEI  485

Qy        479  HKHGMQLRRVG----FN  491
               ::|: |||
Db        486  -RNGIMVRRVGGDNDPN  501
```

```
DR   EMBL; X95098; CAA64475.1; -.
DR   PIR; T06585; T06585.
DR   InterPro; IPR001905; Ammonium_transpt.
DR   Pfam; PF00909; Ammonium_transp; 1.
DR   TIGRFAMs; TIGR00836; amt; 1.
DR   PROSITE; PS01219; AMMONIUM_TRANSP; 1.
KW   Transport; Transmembrane; Multigene family.
FT   TRANSMEM       52     72       POTENTIAL.
FT   TRANSMEM       87    107       POTENTIAL.
FT   TRANSMEM      133    153       POTENTIAL.
FT   TRANSMEM      158    178       POTENTIAL.
FT   TRANSMEM      205    225       POTENTIAL.
FT   TRANSMEM      249    269       POTENTIAL.
FT   TRANSMEM      290    312       POTENTIAL.
FT   TRANSMEM      316    336       POTENTIAL.
FT   TRANSMEM      342    362       POTENTIAL.
FT   TRANSMEM      375    395       POTENTIAL.
```

RESULT 6
AT12_ARATH

```
ID   AT12_ARATH     STANDARD;      PRT;   514 AA.
AC   Q9ZPJ8; Q9SQH8;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Ammonium transporter 1, member 2 (AtAMT1;2).
GN   AMT1.2 OR AT1G64780 OR F13011.9.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RC   STRAIN=cv. C24; TISSUE=Seed;
RX   MEDLINE=9926313; PubMed=10330477;
RA   Gazzarrini S., Lejay L., Gojon A., Ninnemann O., Frommer W.B.,
RA   von Wiren N.;
RT   "Three functional transporters for constitutive, diurnally regulated,
RT   and starvation-induced uptake of ammonium into Arabidopsis roots.";
RL   Plant Cell 11:937-948(1999).
RN   [2]
```

```
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RX   MEDLINE=21016719; PubMed=11130712;
RA   Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA   White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,
RA   Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA   Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA   Dunn P., Etgu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA   Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
RA   Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA   Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA   Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA   Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziali A.,
RA   Militscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA   Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA   Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA   Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA   Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA   Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT   "Sequence and analysis of chromosome 1 of the plant Arabidopsis
RT   thaliana.";
RL   Nature 408:816-820(2000).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RX   MEDLINE=22954850; PubMed=14593172;
RA   Yamada K., Lim J., Dale J.M., Chen H., Shinn P., Palm C.J.,
RA   Southwick A.M., Wu H.C., Kim C.J., Nguyen M., Pham P.K., Cheuk R.F.,
RA   Karlin-Newmann G., Liu S.X., Lam B., Sakano H., Wu T., Yu G.,
RA   Miranda M., Quach H.L., Tripp M., Chang C.H., Lee J.M., Toriumi M.J.,
RA   Chan M.M., Tang C.C., Onodera C.S., Deng J.M., Akiyama K., Ansari Y.,
RA   Arakawa T., Banh J., Banno F., Bowser L., Brooks S.Y., Carninci P.,
RA   Chao Q., Choy N., Enju A., Goldsmith A.D., Gurjal M., Hansen N.F.,
RA   Hayashizaki Y., Johnson-Hopson C., Hsuan V.W., Iida K., Karnes M.,
RA   Khan S., Koesema E., Ishida J., Jiang P.X., Jones T., Kawai J.,
RA   Kamiya A., Meyers C., Nakajima M., Narusaka M., Seki M., Sakurai T.,
RA   Satou M., Tamse R., Vaysberg M., Wallender E.K., Wong C., Yamamura Y.,
RA   Yuan S., Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RT   "Empirical analysis of transcriptional activity in the Arabidopsis
RT   genome.";
RL   Science 302:842-846(2003).
CC   -!- FUNCTION: Involved in high affinity ammonium uptake.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC   -!- TISSUE SPECIFICITY: High expression in root.
CC   -!- SIMILARITY: Belongs to the ammonium transporter (TC 2.A.49)
CC       family.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AF083036; AAD54639.1; -.
DR   EMBL; AF110771; AAD17001; -.
DR   EMBL; AC006193; AAD38253.1; -.
DR   EMBL; AY062571; AAL32649.1; -.
DR   EMBL; AY093374; AAM13373.1; -.
DR   PIR; A96671; A96671.
DR   InterPro; IPR001905; Ammonium_transpt.
DR   Pfam; PF00909; Ammonium_transp; 1.
DR   TIGRFAMs; TIGR00836; amt; 1.
DR   PROSITE; PS01219; AMMONIUM_TRANSP; 1.
KW   Transport; Transmembrane; Multigene family.
FT   TRANSMEM     56     76       POTENTIAL.
FT   TRANSMEM     91    111       POTENTIAL.
FT   TRANSMEM    140    160       POTENTIAL.
FT   TRANSMEM    165    185       POTENTIAL.
FT   TRANSMEM    212    232       POTENTIAL.
FT   TRANSMEM    257    277       POTENTIAL.
```

```
FT   TRANSMEM    291    313       POTENTIAL.
FT   TRANSMEM    328    348       POTENTIAL.
FT   TRANSMEM    351    371       POTENTIAL.
FT   TRANSMEM    380    400       POTENTIAL.
FT   TRANSMEM    431    451       POTENTIAL.
FT   DOMAIN       18     26       POLY-SER.
FT   CONFLICT    198    198       R -> G (IN REF. 1).
FT   CONFLICT    310    310       A -> S (IN REF. 1).
FT   CONFLICT    438    438       V -> I (IN REF. 1).
SQ   SEQUENCE   514 AA;  55013 MW;  86C3C61F044A0CC6 CRC64;
```

```
Query Match             72.7%;  Score 1919; DB 1;  Length 514;
Best Local Similarity   73.7%;  Pred. No. 7.9e-127;
Matches 370;  Conservative  48;  Mismatches  74;  Indels  10;  Gaps  6;

Qy       6  CPAEQLAQLL--GPNTTDASAAASLICGHFAAVDSKFVDTAPAVDNTYLLFSAYLVFSMQ   63
               |  :|:.||| .:|||:  :|  .:  ||| :.   .||||||:||||||||||:|:|
Db       8  CSAVDLSALLSSSSNSTSSLAAATFLCSQISNISNKLSDTTYAVDNTYLLFSAYLVFAMQ   67

Qy      64  LGFAMLCAGSVRAKNTMNIMLTNVLDAAAGGLFYLFGFAFAPGSPSNGFIGKH--FFGL  121
               |||||||||||||||||||||||||||||||||||||||||||| .||.||    ||||
Db      68  LGFAMLCAGSVRAKNTMNIMLTNVLDAAAGAISYYLFGFAFAFGTPSNGFIGRHHSFFAL  127

Qy     122  KDIPS-SSYDYSYFLYQWAFAIAAAGITSGSIAERTQFVAYLIYSSFLTGFVYPVVSHWF  180
               :|:|| ::||||||||:||||||||||||||||||:||||||:||:|||||||||:|||
Db     128  SSYPERPGSDFSFFLYQWAFAIAAAGITSGSIAERTQFVAYLIYSTFLTGFVYPTVSHWF  187

Qy     181  WSPDGWASAFKITDR-LFSTGVIDFAGSGVVHMVGGIAGLWGALIEGPRMGRFDHAGRAV  239
               ||||||||||||  | .|   |:|.||:|||||||||||||.:|: .|:.|||:|.::||
Db     188  WSSDGWASASRSDNNLLFGSGAIDFAGSGVVHMVGGIAGLCGALVEGPRIGRFDRSGRSV  247

Qy     240  ALRGHSASLVVLGTFLLMFGWYGFNPGSFNKILLTYGNSGNYYGQWSAVGRTAVTTTLAG  299
               ||||||||||||||||||||||||||||||| |||||||:||||||||||||||||||:|
Db     248  ALRGHSASLVVLGTFLLMFGWYGFNPGSFLTILKGYDKSRPYYGQWSAVGRTAVTTTLSG  307

Qy     300  STAALTTLFGKRVISGHWNVTDVCNGLLGGFAAITAGCSVVEPWAAIVCGFVASIVLIAC  359
               .||||||||| | ::|||||:||||||||||||||||.|:||||||||||:||::|| ||
Db     308  CTAALTTLFSKRLLAGHWNVIDVCNGLLGGFAAITSGCAVVEPWAAIVCGFVASWVLIGF  367

Qy     360  NKLAEKVKFDDPLEAAQLHGGCGTWGVIFTALFAKKEYVKEVYGLGRAHGLLMGGGGKLL  419
               |||:|||.|||||:||||||||.||.|:.:|||||||||||:| .:||||.||:|||||:
Db     368  NLLAKKLKYDDPLEAAQLHGGCGAWGLIFTGLFARKEYVNEIYSGDRPYGLFMGGGGKLL  427

Qy     420  AAHVIQILVIAGWVSATMGPLFWGLNKLKLRISSEDELAGMDMTRHGGFAYAYEDDE--  477
               |||||||:|||||:||||||:|||||:| |::|||:||||||||:|||:||||  |.
Db     428  AAQIVQIIVIVGWVTVTMGPLFYGLHKMNLLRISAEDEMAGMDMTRHGGFAYAYNDEDDV  487

Qy     478  -THKHGMQLRRVGPNA-STPT  497
               .||:||:|||: |||. |.||
Db     488  STKPWGHFAGRVEPTSRSSTPT  509
```

```
RESULT 7
AT13_LYCES
ID   AT13_LYCES             STANDARD;      PRT;    460 AA.
AC   Q9FVN0;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Ammonium transporter 1, member 3 (LeAMT1;3).
GN   AMT1.3.
OS   Lycopersicon esculentum (Tomato).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC   lamiids; Solanales; Solanaceae; Solanum.
OX   NCBI_TaxID=4081;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Leaf;
RX   MEDLINE=20208060; PubMed=10743657;
RA   von Wiren N., Lauter F.-R., Ninnemann O., Gillissen B., Walch-Liu P.,
RA   Engels C., Jost W., Frommer W.B.;
RT   "Differential regulation of three functional ammonium transporter
```

```
RT   genes by nitrogen in root hairs and by light in leaves of tomato.";
RL   Plant J. 21:167-175(2000).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC   -!- TISSUE SPECIFICITY: Leaves.
CC   -!- INDUCTION: Highest expression in darkness.
CC   -!- SIMILARITY: Belongs to the ammonium transporter (TC 2.A.49)
CC       family.
CC   ------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ------------------------------------------------------------------------
DR   EMBL; AF118858; AAG11397.1; -.
DR   InterPro; IPR001905; Ammonium_transpt.
DR   Pfam; PF00909; Ammonium_transp; 1.
DR   TIGRFAMs; TIGR00836; amt; 1.
DR   PROSITE; PS01219; AMMONIUM_TRANSP; 1.
KW   Transport; Transmembrane; Multigene family.
FT   TRANSMEM    15     37       POTENTIAL.
FT   TRANSMEM    50     72       POTENTIAL.
FT   TRANSMEM    98    117       POTENTIAL.
FT   TRANSMEM   124    146       POTENTIAL.
FT   TRANSMEM   166    188       POTENTIAL.
FT   TRANSMEM   209    227       POTENTIAL.
FT   TRANSMEM   255    277       POTENTIAL.
FT   TRANSMEM   305    327       POTENTIAL.
FT   TRANSMEM   337    356       POTENTIAL.
FT   TRANSMEM   377    399       POTENTIAL.
SQ   SEQUENCE   460 AA;   49652 MW;   6D9CFC6C2C2395BA CRC64;

Query Match             67.2%;  Score 1591;  DB 1;  Length 460;
Best Local Similarity   60.3%;  Pred. No. 5.5e-104;
Matches  301; Conservative   56; Mismatches   81; Indels   10; Gaps    5;

Qy   46 AVDNTYLLFSAYLVFSMQLGFAMLCAGSVRAKNTMNIMLTNVLDAAAGGLFYYLFGFAFA 105
         :: :   |||||||||.|||||||||||||||:.|||||||||.. ||.|.||||||||
Db   12 SINAIYLLFSAYLVFVMQLGFAMLCAGSVRAKNAMNIMLTNVDAVVGSLSYYLFGFAFA 71

Qy  106 FGSPSNGFIGKHFFGLKDIPSSSYDYSYFLYQWAFAIAAAGITSGSIAERTQFVAYLIYS 165
         ||. ||.||  .|.|||||||||||||.|||||||||..|||||||||||||||.||..|
Db   72 FGD-SNPFIGASYFALKDIPSSSYDYSFFLYQWAFAIAVAGITSGSIAERTQFFAYLVFS 130

Qy  166 SFLTGFVYPVVSHWFWSPDGWASAFKITDRLFSTGVIDFAGSGVVHMVGGIAGLWGALIE 225
         .||||||||||.||.|| .||.|  .   .:   .:.:.||||||||.||||||.|.::|
Db  131 FFLTGFVYPVVAHWLWSSNGWLSP-NSTYLLFGSSAIDFAGSGVVHLVGGIAGFWGSIVE 189

Qy  226 GPRMGRFDHAGRAVALRGHSASLVVLGTFLLWFGWYGFNPGSFNKILLTYGNSGNYYGQW 285
         ||| |||:  | ||.|||.|.:||||||.|||||||.|||||||||:..|.: .. .|.|
Db  190 GPRVGRFDAFGNFVKMRGHNATLVVLGTLLLWFGWFGFNPGSFNKILVAYPHMAD-QGNW 248

Qy  286 SAVGRTAVTTLLAGSTAALTTLFGKRVISGHWNVTDVCNGLLGGFAAITAGCSVVEPWAA 345
         :.|||||||||.||||| :.|||| |.: |||:. |||||:||||.|||:||||||||||
Db  249 TSVGRTAVTTLLAGSTAGIVTLFGRRLLVGHWDAMDVCNGVLGGFVAITSGCSVVEPWAA 308

Qy  346 IVCGFVASIVLIACNKLAEKVKFDDPLEAAQLHGGCGTWGVIFTALFAKKEYVKEVYGLG 405
         |.|||.|.::. .. |:::||.|.|||| ||||||||.||.|||.||||.|.| ..| ..
Db  309 ILCGFCAAWLIGLNIILALKFKFDDPLERAQLHGGCGAWGLIFTGLFAKEEFVLQAYNSG 368

Qy  406 ----RAHGLLMGGGGKLLAAHVIQILVIAGWVSATMGPLFWGLNKLKLLRISSEDELAG 460
             :. ..:||||  | ||  .:::..: .||||.|||||| | .|:.||||:||:||
Db  369 KTQIIRPSGLILGGGWGLFGAQIVELLSIVVWVSLTMGPLFYLLQKLGILRISSDEEVAG 428

Qy  461 MDMTRHGGFAY--AYEDDETHKGMQLR 486
         :|:..|||:||   :::: . ..| .||
Db  429 LDISSHGGYAYDASQEESNARFYGEYLR 456

RESULT 8
AMT1_CAEEL
```

```
ID   AMT1_CAEEL     STANDARD;      PRT;   534 AA.
AC   P54145;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Putative ammonium transporter C05E11.4.
GN   C05E11.4.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RA   Geisel C.;
RL   Submitted (APR-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Involved in the uptake of ammonia (Probable).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC   -!- SIMILARITY: Belongs to the ammonium transporter (TC 2.A.49)
CC       family.
CC   ------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ------------------------------------------------------------------------
DR   EMBL; U53338; AAA96191.1; -.
DR   PIR; T15414; T15414.
DR   WormPep; C05E11.4; CE06770.
DR   InterPro; IPR001905; Ammonium_transpt.
DR   InterPro; IPR008960; CBD9-like.
DR   Pfam; PF00909; Ammonium_transp; 1.
DR   TIGRFAMs; TIGR00836; amt; 1.
DR   PROSITE; PS01219; AMMONIUM_TRANSP; 1.
KW   Hypothetical protein; Transport; Transmembrane.
FT   TRANSMEM    31     51       POTENTIAL.
FT   TRANSMEM    69     89       POTENTIAL.
FT   TRANSMEM   115    135       POTENTIAL.
FT   TRANSMEM   139    159       POTENTIAL.
FT   TRANSMEM   184    204       POTENTIAL.
FT   TRANSMEM   223    243       POTENTIAL.
FT   TRANSMEM   263    283       POTENTIAL.
FT   TRANSMEM   291    311       POTENTIAL.
FT   TRANSMEM   318    338       POTENTIAL.
FT   TRANSMEM   346    366       POTENTIAL.
FT   TRANSMEM   401    421       POTENTIAL.
SQ   SEQUENCE   534 AA;   58490 MW;   C17CFE44E0DC3D05 CRC64;

Query Match             26.8%;  Score 706;  DB 1;  Length 534;
Best Local Similarity   36.0%;  Pred. No. 3.7e-42;
Matches  164; Conservative   74; Mismatches  185; Indels   32; Gaps    9;

Qy   41 VDTAF--AVDNTYLLFSAYLVFSMQLGFAMLCAGSVRAKNTMNIMLTNVLDAAAGGLPYY 98
         :..     :::..:||||||:||:|||:|||:|||||||||||||||:||::::::::
Db   21 LETGFYENVNSFFLCSMALIIFFMQCGFAYLEAGAVRSKNTTNILIKNLLDSCICIIGYW 80

Qy   99 LFGFAFAPGSPSNG--FIGKHFPGLKDIPSSSYDYSYFLYQWAFAIAAAGITSGSIAER 155
         :: |.|.        :..||  ||| |::|   ..   .||   ...  | ...
Db   81 AIGWALAYGDSGEGVNLFVGHSQFFL---SGFSDYPRFFFQYVFSATAATIVSGAVAER 136

Qy  156 TQFVAYLIYSSFLTGFVYPVVSHWFWSPDGWASAFKITDRLFSTGVIDFAGSGVVHMVGG 215
         |||: |||.:||||||||||.:||.|   .     ...   |||.|||..||||||.||
Db  137 CEFITVYTYCTVISTFIYPVLTHWGWTENGW-MAKGITSGIIDTKYDDFAGSGLVHLCGG 195

Qy  216 IAGLWGALIEGPRMGRF--DHAGRAVALRGHSASLVVLGTFLLWFGWYGFNPGSFNKILL 273
         ||.||.|:||||:||:      :| .:||||.:|:||:|..|.||||:||.||.:|:
Db  196 SISFLAAWIMGPRIGKFPDDEDDESDEILGHSVPFTALGGFILMFGFLAFNGGSVASI-- 253

Qy  274 TYGNSGNYYGQWSAVGRTAVTTLLGKRVISGHWNVTDVCNGLLGGFAAI 333
         |:  ...  .        ...|.||..||||| : ||| |||  ||||
```

```
Db  254 ------SHAGDGHTVALAMINTILSGAFAALIYLGVHYYQHGKWTLLLTINACLSGMVAA 307
Qy  334 TAGCSVVEPWAAIVCGFVASIVLIACNKLABKVKFDDPLEAAQLHGGCGTWGVIFTALFA 393
        ||| :||||           ::| || |||      || :|: || |:  ||||||||
Db  308 CAGCNKMEPWACIWVGLGAGLIYLAFSKLMIRLKIDDPLDAPAVHAGGGFWGLMSSSSIIS 367
Qy  394 KKEYVKEVYGLGRAHGLLMGGGGKLLAA------HVIQILVIAGWVSATMGPLFWGLNKL 447
        :|              |||||||||||       |  :| :| |||
Db  368 --HGGVAYALADAVSGAKNSGDHLTQAFAQIGWQMICALALIIAWSLGVMLPIFWILKKT 424
Qy  448 KLLRISSEDELAGMDMTRHGGFAYAYEDDETHKHG 482
        |  : |:|: || |||| ||||||| :: || ||
Db  425 GKLRVSEEVINGLDVFKHGEMAYPL--RAYGHG 456
```

```
RESULT 9
Y108_SYNY3
ID   Y108_SYNY3     STANDARD;      PRT;      507 AA.
AC   P54147;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Putative ammonium transporter sll0108.
GN   SLL0108.
OS   Synechocystis sp. (strain PCC 6803).
OC   Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX   NCBI_TaxID=1148;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96127529; PubMed=8590279;
RA   Kaneko T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,
RA   Sugiura M., Tabata S.;
RT   "Sequence analysis of the genome of the unicellular cyanobacterium
RT   Synechocystis sp. strain PCC6803. I. Sequence features in the 1 Mb
RT   region from map positions 64% to 92% of the genome.";
RL   DNA Res. 2:153-166(1995).
```
```
DR   EMBL; D64004; BAA10631.1; -.
DR   PIR; S76687; S76687.
DR   InterPro; IPR001905; Ammonium_transpt.
DR   Pfam; PF00909; Ammonium_transp; 1.
DR   TIGRFAMs; TIGR00836; amt; 1.
DR   PROSITE; PS01219; AMMONIUM_TRANSP; 1.
KW   Hypothetical protein; Transport; Transmembrane; Complete proteome.
FT   TRANSMEM     44     64       POTENTIAL.
FT   TRANSMEM     90    110       POTENTIAL.
FT   TRANSMEM    128    148       POTENTIAL.
FT   TRANSMEM    154    174       POTENTIAL.
FT   TRANSMEM    178    198       POTENTIAL.
FT   TRANSMEM    211    231       POTENTIAL.
FT   TRANSMEM    245    265       POTENTIAL.
FT   TRANSMEM    286    306       POTENTIAL.
FT   TRANSMEM    315    335       POTENTIAL.
FT   TRANSMEM    347    367       POTENTIAL.
FT   TRANSMEM    370    390       POTENTIAL.
FT   TRANSMEM    399    419       POTENTIAL.
FT   TRANSMEM    438    458       POTENTIAL.
SQ   SEQUENCE    507 AA;  53580 MW;  1840AB99FFA08E6A CRC64;
```

```
Query Match             26.6%;   Score 703;  DB 1;  Length 507;
Best Local Similarity   36.2%;   Pred. No. 5.7e-42;
Matches 160;  Conservative  63;  Mismatches 173;  Indels  46;  Gaps  8;
```

```
Qy   51 YLLFSAYLVFSMQLGFAMLCAGSVRAQNTMNILMLTNVLDAAAGGLFYLFGFAFAFGSPS 110
        :||  :   |::|||||||||||   ::  ||||:|||| | :|| |||:|||||||||
Db   92 WLIAAILVIFNNAGFGMLETGLCRQKNAVNIIJTKNLIVFALATIAYWAIGFSLMFGSSG 151
Qy  111 NGFI------GKHF-FGLKDIPSSYDYSYPLYQWAFAIAAAGITSGSIAERTQFVAY 161
        :|         ||: | : ::||
Db  152 NPFVGFGGFFLSGDHTNYGLSPFPEGLPVAVFFLPQVAFSATAATIVSGAVAERIKFNEF 211
Qy  162 LIYSSFLTGFVYPVVSHWFWSPDGWASAFKITDRLFSTGVIDFAGSGVVHMVGGIAGLWG 221
        ||:| |: |||:  || |:| |||          |||||| |||||:|||||| | ||
Db  212 LIFSVLLVGIAYPITGHWVWDAGGW---------LYTMGFMDFAGSTVVHSVGGWAALAG 262
Qy  222 ALIEGPRMGRFDHAGRAVALRGHSASLVVLGTFLLWFGWYGFNPGSFNKILLTYGNSGNY 281
        |::||||:|:| : ||     ||  :  ::| ::||:||:|||||||          Q 308
Db  263 AFLLGPRLGKFVD-GRPGAIPGHNMGFAMLGCLILWIGWFGFNPGS----------Q 308
Qy  282 YGQWSAVGRTAVTTLLAGSTAALTTLFGKRVISGHWNVTDVCNGLLGGFAAITAGGSVVE 341
        |   :||  ||  | |:|::||:| |  | |||:||||      ::||| ::| ||:|
Db  309 LAADQACAYIAVTTNLAASAGGLTATFTSWLKDGKPDLTMVINGVLAGLVGITAGCAGVS 368
Qy  342 PWAAIVCGFVASLVLIACNKLABKVKFDDPLEAAQLHGGCGTWGVIFTALFAKKEVVKEV 401
         | ||| | |  :|:  : |::    || ||  : :|  |  :|  ::|
Db  369 YWGSVIIGGIAGILVVYSVAFFDKIKIDDPVGAISVHLVNGVWGTLAVGFFNMEK----- 423
Qy  402 YGLGRAHGLLMGGGGKLLAAHVIQILVIAGWVSATMGPLFWGLNKLKL-LRISSEDELAG 460
        ||| |  ||:||:    |    | :|:|| :      |||| |:|| |  |: || ::
Db  424 --------GLFYGGGINQLIIQIVGILAI-GAFTAIFSFVVWAAILKQTMGIRVSGEEMIG 475
Qy  461 MDMTRHGGFAYAYEDDETHKHG 482
        :|: :||  :||    :|: |
Db  476 LDIGEHGMEAYTGFVKETDSFG 497
```

```
RESULT 10
YA17_SYNY3
ID   YA17_SYNY3     STANDARD;      PRT;      442 AA.
AC   P72935;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Putative ammonium transporter sll1017.
GN   SLL1017.
OS   Synechocystis sp. (strain PCC 6803).
OC   Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX   NCBI_TaxID=1148;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97061201; PubMed=8905231;
RA   Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA   Miyajima N., Hirosawa M., Sugiura M., Sasamoto S., Kimura T.,
RA   Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K.,
RA   Okumura S., Shimpo S., Takeuchi C., Wada T., Watanabe A.,
RA   Yamada M., Yasuda M., Tabata S.;
RT   "Sequence analysis of the genome of the unicellular cyanobacterium
RT   Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT   entire genome and assignment of potential protein-coding regions.";
RL   DNA Res. 3:109-136(1996).
```
```
DR   EMBL; D90901; BAA16952.1; -.
DR   PIR; S74801; S74801.
DR   InterPro; IPR001905; Ammonium_transpt.
```

```
DR   InterPro; IPR002229; RhesusRHD.
DR   Pfam; PF00909; Ammonium_transp; 1.
DR   PRINTS; PR00342; RHESUSRHD.
DR   TIGRFAMs; TIGR00836; amt; 1.
DR   PROSITE; PS01219; AMMONIUM_TRANSP; 1.
KW   Hypothetical protein; Transport; Transmembrane; Complete proteome.
FT   TRANSMEM      5     25       POTENTIAL.
FT   TRANSMEM     44     64       POTENTIAL.
FT   TRANSMEM     81    101       POTENTIAL.
FT   TRANSMEM    104    124       POTENTIAL.
FT   TRANSMEM    133    153       POTENTIAL.
FT   TRANSMEM    155    175       POTENTIAL.
FT   TRANSMEM    193    213       POTENTIAL.
FT   TRANSMEM    240    260       POTENTIAL.
FT   TRANSMEM    269    289       POTENTIAL.
FT   TRANSMEM    299    319       POTENTIAL.
FT   TRANSMEM    325    345       POTENTIAL.
FT   TRANSMEM    354    374       POTENTIAL.
FT   TRANSMEM    386    406       POTENTIAL.
SQ   SEQUENCE    442 AA;  47172 MW;  F3683BF23D96539B CRC64;

Query Match             25.7%;  Score 678.5;  DB 1;  Length 442;
Best Local Similarity   35.9%;  Pred. No. 2.5e-40;
Matches  168;  Conservative   61;  Mismatches 152;  Indels  87;  Gaps  15;

Qy    38 SKFVDTAFAVDNTYLLFSAYLVFSMQLGFAMLCAGSVRAKNTMNIMLTNVLDAAAGGLFY 97
          : :|:|||: :||:||| ||| |||| |||
Db    32 TSIAEVTYAINNLFLLAAAVLVLFMQAGFAMLEAGLSSHKNTVNVLFKNTFDVCGVLLY 91

Qy    98 YLFGFAFAPGSPS--NGFIGKHFFG----LKDIPSSSYDYSYFLYQWAFAIAAAGITSGS 151
          :||| || |||    ||||:||
Db    92 FLFGYISLMYGENPVLGGFFGWGGFGITNNLDNVEGLSPQVDW-LFQAARAATAIVSGA 150

Qy   152 IAERTQFVAYLLYSSFLTGFVYPVVSWFWSPDGWASAFKITDRLFSTGVIDFAGSGVVH 211
          : |||||| ||||     ||||| |||
Db   151 VMGRMYFKAYLLYSAVITGLVIPISGHWKG-GGW------LDKL---GFHDFAGSLLVH 200

Qy   212 MVGGIAGLWGALIEGPRMGRFDHAGRAVALRGH---SASLVVLGTFLLWFGWYGFNPGS 267
          ||| || :: ||||:||||    :|||
Db   201 SVGGFAALAAVVMGPRIGRFE--GNKINSLGYQGITSSS---LGVFILWGWYGFNPGS 255

Qy   268 FNKILLTYGNSGNYYGQWSAVGR-------TAVTTTLAGSTRAALTTLFGKRVISGHW-- 317
          ||                              |:
Db   256 -----------------QLAFVGALNTNTTMLIAVNTTLSAAAGGLAAL-----AFDWIT 293

Qy   318 ----NVTDVCNGLLGGFAAITAGCSVVEPWAAIVCGFVASIVLIACNKLAEKVKFDDP 371
          :     ||||||||||    :||:
Db   294 ENKRKPNLLVTLNGILGLGLVGITTAGCDTVSNWSAIAIGVVAGILSVLGTKLLDRLRIDDG 353

Qy   372 LEAAQLHGGCGTWGVIFTALFAKKEYVKSVYGLGRAHGLLMGGGGKLLAAHVIQILVIAG 431
          : |||||||| |||     :||
Db   354 VGAWPVEGLCGIWGGIAVGIFSTNVEHK-----------------LSAQIVGSLVIPF 394

Qy   432 WVSATMGPLFWGLNKLKLKLRISSEDELAGMDMTRHG----GFAYAYED 475
          ||     :|||:       : ||
Db   395 WAFITMFFLFYVMDLWGTLRVKPSQEKVGLDIVEHGQTEKGVEIAFED 442


RESULT 11
AMT3_CAEEL
ID   AMT3_CAEEL    STANDARD;    PRT;   687 AA.
AC   Q21565;
DT   15-JUL-1998 (Rel. 36, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Putative ammonium transporter amt-3.
GN   AMT-3 OR M195.3.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
```

```
RC   STRAIN=Bristol N2;
RA   Burton J.;
RL   Submitted (OCT-1995) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   REVISIONS.
RA   Durbin R.;
RL   Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Involved in the uptake of ammonia (Probable).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC   -!- SIMILARITY: Belongs to the ammonium transporter (TC 2.A.49)
CC       family.
CC   --------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------
DR   EMBL; Z66498; CAA91293.2; -.
DR   PIR; T23804; T23804.
DR   WormPep; M195.3; CE32912.
DR   InterPro; IPR001905; Ammonium_transpt.
DR   Pfam; PF00909; Ammonium_transp; 1.
DR   TIGRFAMs; TIGR00836; amt; 1.
DR   PROSITE; PS01219; AMMONIUM_TRANSP; 1.
KW   Hypothetical protein; Transport; Transmembrane.
FT   TRANSMEM     39     59       POTENTIAL.
FT   TRANSMEM     77     97       POTENTIAL.
FT   TRANSMEM    134    154       POTENTIAL.
FT   TRANSMEM    162    182       POTENTIAL.
FT   TRANSMEM    196    216       POTENTIAL.
FT   TRANSMEM    240    260       POTENTIAL.
FT   TRANSMEM    272    292       POTENTIAL.
FT   TRANSMEM    299    319       POTENTIAL.
FT   TRANSMEM    323    343       POTENTIAL.
FT   TRANSMEM    352    372       POTENTIAL.
FT   TRANSMEM    404    424       POTENTIAL.
SQ   SEQUENCE    687 AA;  73937 MW;  90EA97A92891BE01 CRC64;

Query Match             22.0%;  Score 579.5;  DB 1;  Length 687;
Best Local Similarity   31.8%;  Pred. No. 3.2e-33;
Matches  156;  Conservative   86;  Mismatches 169;  Indels  79;  Gaps  19;

Qy    14 LLGPNTT--DASARAASLICHFAAVD-SKFVDTAFAVDNTYLLFSAYLVFSMQLGFAMLC 70
          :  :||:   |:||              ||||||||||||| :|:|:||:||||||:
Db     1 MAGPEGSIFNASAMQIVQIHHYAEGSVTPEVDKLYQDDAVWIISSFIIFTMHSGFGLLE 60

Qy    71 AGSVRAKNTMNIMLTNVLDAAAGGLFYLFGFAFAFG------SPSNGFIGKHF------ 118
          :||| ||:|||| ||:|||||||||||||||||||       :||||||||
Db    61 SGSVSAKDEVNIMVKNVVDVVFGGLSYWSCGFGFSYGDIPEWRNPVYVGF-GKFPFYDPTRD 119

Qy   119 FGLKD-IPSSSYDYSYFLYQWAFAIAAAGITSGSIAERTQFVAYLLYSSFLTGFVYPVVS 177
          |||  |  : :||  :||||||:||||||  ||| || ||||| |||  ||:
Db   120 YGTRETINQEGWSYASFLFQLSLATTASTIVSGAVAERAKLKSYILLGCIVI-LIQAlPA 178

Qy   178 HWFWSPDGWASAFKITDRLFSTGVIDFAGSGVVHMVGGIAGLWGGIAGLWGALIEGPRMGRFDHAGR 237
          || :: |:                 |:||||| |||:|||| ||   |||| :|||
Db   179 HWWDKEG----------VFYKKGVVDFAGCSAVHLVGGIIGLIATVFLKPRRNRFNEDS- 228

Qy   238 AVALRGHSASLVVLGTFLLWFGWYGFNPGSFNKILLTYGNSGNYYGQWSAVGRTAVTTTL 297
          |: :||:|     ::||||||||||| |:       :|| | ||
Db   229 --VHQMSSFNALLGTFLLWWGWFGINAGS------VWGITG--GRWRLGARAAVATIM 277

Qy   298 A---GSTAALTTLFGKRVISGHWNVTDVCNGLLGGFAAITAGCSVVEPWAAIVCGFVASI 354
          |           |:  |||     :::     :|:|    :| :
Db   278 ASIGGATAITISFVK--TKKLQVNFLINGILSSIVSITAICAVSRPWHALVIGSISSV 334

Qy   355 VLIACNKLAEKVKFDDPLEAAQLHGGCGTWGVIFTALFAKKEYVKEVYGLGRA---HGL 410
          ||:| ||:  :|:: :|:|:|   |||    |||:| |  ::: : |:|:
Db   335 FSIAVLPLLDRLHIDDPVGIVPIHLTSSIWGMIAVGIFCEEDKY------LGSATNNRSGL 389
```

```
Qy    411 LMGGGGKLLAAHVIQILVIAGWV----------------SATMGPL-FWGLNKLKL-LRISSEDE 457
              |||:|      |:|             ||||||| |||:||: :|
Db    390 LYSWSFELL--------------------WVQLQCTAAILIYSATTGFLALFLISKSPLGLRVTDYEE 437

Qy    458 LAGMDMTRHG 467
              |:|||
Db    438 QIGADVIEHG 447

RESULT 12
/537_SYNY3
ID   Y537_SYNY3           STANDARD;       PRT;    541 AA.
AC   P54148;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Putative ammonium transporter sll0537.
GN   SLL0537.
OS   Synechocystis sp. (strain PCC 6803).
OC   Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX   NCBI_TaxID=1148;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96127529; PubMed=8590279;
RA   Kaneko T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,
RA   Sugiura M., Tabata S.;
RT   "Sequence analysis of the genome of the unicellular cyanobacterium
RT   Synechocystis sp. strain PCC6803. I. Sequence features in the 1 Mb
RT   region from map positions 64% to 92% of the genome.";
RL   DNA Res. 2:153-166(1995).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC   -!- SIMILARITY: Belongs to the ammonium transporter (TC 2.A.49)
CC       family.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; D64006; BAA10864.1; -.
DR   PIR; S76017; S76017.
DR   InterPro; IPR001905; Ammonium_transpt.
DR   InterPro; IPR003660; HAMP.
DR   Pfam; PF00909; Ammonium_transp; 1.
DR   SMART; SM00304; HAMP; 1.
DR   TIGRFAMs; TIGR00836; amt; 1.
DR   PROSITE; PS01219; AMMONIUM_TRANSP; 1.
KW   Hypothetical protein; Transport; Transmembrane; Complete proteome.
FT   TRANSMEM      4     26       POTENTIAL.
FT   TRANSMEM     44     64       POTENTIAL.
FT   TRANSMEM     86    106       POTENTIAL.
FT   TRANSMEM    117    137       POTENTIAL.
FT   TRANSMEM    161    181       POTENTIAL.
FT   TRANSMEM    203    223       POTENTIAL.
FT   TRANSMEM    235    255       POTENTIAL.
FT   TRANSMEM    260    280       POTENTIAL.
FT   TRANSMEM    283    303       POTENTIAL.
FT   TRANSMEM    316    336       POTENTIAL.
FT   TRANSMEM    356    376       POTENTIAL.
SQ   SEQUENCE   541 AA;  58129 MW;  A7A0A5CB2FD2475 CRC64;

Query Match             21.6%;  Score 569.5;  DB 1;  Length 541;
Best Local Similarity   33.2%;  Pred. No. 1.3e-32;
Matches  153;  Conservative  58;  Mismatches 191;  Indels  59;  Gaps  12;

Qy     46 AVDNTYLLFSAYLVFSMQLGFAMLCAGSVRAKNTMNIMLTNVLDAAAGGLFYYLFGFAFA 105
              ||   :|||| |||:|||:||:|||::| |:||||:|:::: :: : ||||||
Db      3 SIDTLWLLLCAGLVFFMQAGFMCLESGLTRSKNSINVAIKNFADFGISVALFWSFGFSIM 62

Qy    106 FGSPSNGFIGKHFFGLKDIPSSSYDYSYFLYQWAFAIAAGITSGSIAERTQFVAVLIYS 165
```

```
Db      353 LVGTLLVGLLAAPE-APAINGVAGVSKGLFYGGGFAQLERQALGACSVLVYSGIITLILA 411

Qy      439 PLFWGLNKLKL---LRISSEDELAGMDMTRH--GGFAYAYEDD--------ETHKHGMQ 484

Db      412 LI------LKFTIGLRLDAEQESTGIDEAEHAESGYDFAVASGSVLPPRVTVEDSRNGIQ 465

Qy      485 LRRVGPNASSTP 496
               |:|:||:||
Db      466 -ERIGQKVEAEP 476
```

RESULT 14
NRGA_BACSU
ID   NRGA_BACSU   STANDARD;   PRT;   404 AA.
AC   Q07429;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Probable ammonium transporter (Membrane protein nrgA).
GN   NRGA OR BSU36510.
OS   Bacillus subtilis.
OC   Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX   NCBI_TaxID=1423;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94110214; PubMed=8282685;
RA   Wray L.V. Jr., Atkinson M.R., Fisher S.H.;
RT   "The nitrogen-regulated Bacillus subtilis nrgAB operon encodes a
RT   membrane protein and a protein highly similar to the Escherichia coli
RT   glnB-encoded PII protein.";
RL   J. Bacteriol. 176:108-114(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168;
RX   MEDLINE=98015417; PubMed=9353933;
RA   Presecan E., Moszer I., Boursier L., Cruz Ramos H.C., De La Fuente V.,
RA   Hullo M.-F., Lelong C., Schleich S., Sekowska A., Song B.H.,
RA   Villani G., Kunst F., Danchin A., Glaser P.;
RT   "The Bacillus subtilis genome from gerBC (311 degrees) to licR (334
RT   degrees).";
RL   Microbiology 143:3313-3328(1997).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168;
RX   MEDLINE=98044033; PubMed=9384377;
RA   Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA   Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA   Borriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA   Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA   Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA   Denizot F., Devine K.M., Duesterhoft A., Ehrlich S.D., Emmerson P.T.,
RA   Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA   Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA   Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA   Guiseppi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
RA   Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA   Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA   Kobayashi Y., Koetter P., Koningstein G., Krogh S., Kumano M.,
RA   Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA   Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA   Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA   Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA   Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA   Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA   Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
RA   Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA   Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
RA   Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA   Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA   Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA   Viari A., Wambutt R., Wedler E., Wedler H., Weitzenegger T.,
RA   Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA   Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;

RA   Garnier T., Eiglmeier K., Camus J.-C., Medina N., Mansoor H.,
RA   Pryor M., Duthoy S., Grondin S., Lacroix C., Monsempe C., Simon S.,
RA   Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA   Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RT   "The complete genome sequence of Mycobacterium bovis.";
RL   Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
CC   -!- FUNCTION: Involved in the uptake of ammonia (Probable).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC   -!- SIMILARITY: Belongs to the ammonium transporter (TC 2.A.49)
CC   family.
CC   --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --
DR   EMBL; Z74697; CAA98980.1; -.
DR   EMBL; AE007120; AAK47314.1; -.
DR   EMBL; BX248344; CAD96631.1; -.
DR   PIR; H70747; H70747.
DR   TIGR; MT2988; -.
DR   TuberculList; Rv2920c; -.
DR   InterPro; IPR001905; Ammonium_transpt.
DR   Pfam; PF00909; Ammonium_transp; 1.
DR   TIGRFAMs; TIGR00836; amt; 1.
DR   PROSITE; PS01219; AMMONIUM_TRANSP; 1.
KW   Transport; Transmembrane; Complete proteome.
FT   TRANSMEM   16   36   POTENTIAL.
FT   TRANSMEM   53   73   POTENTIAL.
FT   TRANSMEM   121   141   POTENTIAL.
FT   TRANSMEM   150   170   POTENTIAL.
FT   TRANSMEM   192   212   POTENTIAL.
FT   TRANSMEM   229   249   POTENTIAL.
FT   TRANSMEM   257   277   POTENTIAL.
FT   TRANSMEM   290   310   POTENTIAL.
FT   TRANSMEM   312   332   POTENTIAL.
FT   TRANSMEM   344   364   POTENTIAL.
FT   TRANSMEM   394   414   POTENTIAL.
SQ   SEQUENCE   477 AA; 49241 MW; 97CB46A1C1063DFB CRC64;

Query Match   18.8%; Score 496; DB 1; Length 477;
Best Local Similarity   29.7%; Pred. No. 1.5e-27;
Matches 146; Conservative 79; Mismatches 195; Indels 72; Gaps 17;

Qy   48 DNTYLLFSAYLVPSMQLGFAMLCAGSVRAKNTMNIMLTNVLDAAAGGLFYYLFGFAFAFG 107
               :||:||| ||| ||::|: ||:||:: : :|:|:| ||
Db   14 DTAWMLVSSALVLLMTPGLAFFYCGMVRSKSVLNMIMSISAMGVVTVLWALYGYSIAFG 73

Qy   108 SPSNGFIG--KHFFGLK---------------DIP-SSSYDYSYFL-YQWAPAIAA 144
               :|:||||:|
Db   74 DDVGNIAGNPSQYWGLKGLIGVNAVAADPSTQTAAVNIPLAGTLPATVFVAFQLMFAIIT 133

Qy   145 AGITSGSSIAERTQFVAYLLYSSFLTGFVYPVSHWFWSPDGWASAFK--ITDRLFSTGVI 202
               |||:|:|:| ||::||:: |:|
Db   134 VALISGAVADRLKFGAWLLFAGLWATFVYFPPVAHWVFAFDGFAAEBRGGWIANKLHA---I 190

Qy   203 DFAGSGVVHMVGGIAGLWGALIEGPRMGRFDHAGRAVALRGHSASLVVLGTFLIWFGWYG 262
               ||||:|||||
Db   191 DFAGGTAVHINAGVAALMLAIVLGKRRG-----WPATLFRPHNLPFVMLGAALLNFGWYG 245

Qy   263 FNPGSFNKILLTYGNSGNYYGQWSAVGRTAVTTLAGSTAALTTLFGKRVISGHWNVTDV 322
               |:
Db   246 FNAGS-----ATTAN------GVAGATFVTTTIATAAAMLGWLLTERVRDGKATTLGA 292

Qy   323 CNGLLGGFAAITAGCSVVEPWAAIVCGFVASIVLIACNKLAEKVKFDDPLEAAQLHGGCG 382
               |:||:| :||:||::
Db   293 ASGIVAGLVAITPSCSSVNVLGALAVGVSAGVLCALAVGLKFKLGFDDSLDVVGVHLVGG 352

Qy   383 TWGVIFTALFAKKEYVKEVYGL-GRAHGLLMGGGGKLL--AAHVIQILVIAGWVSATMG 438
               | :||||||||: || |||||||| ||: : :
```

```
RT   "The complete genome sequence of the Gram-positive bacterium Bacillus
RT   subtilis";
RL   Nature 390:249-256(1997).
CC   -|- FUNCTION: May participate in nitrate utilization or facilitate
CC       adaptation to growth on this medium. NrgA may function as an
CC       ammonium transport protein.
CC   -|- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC   -|- SIMILARITY: Belongs to the ammonium transporter (TC 2.A.49)
CC       family.
CC   ----------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
DR   EMBL; L03216; AAA17399.1; -.
DR   EMBL; Z82987; CAB05374.1; -.
DR   EMBL; Z99122; CAB15668.1; -.
DR   PIR; A36865; A36865.
DR   SubtiList; BG10869; nrgA.
DR   InterPro; IPR001905; Ammonium_transpt.
DR   Pfam; PF00909; Ammonium_transp; 1.
DR   TIGRFAMs; TIGR00836; amt; 1.
DR   PROSITE; PS01219; AMMONIUM_TRANSP; 1.
KW   Transport; Transmembrane; Complete proteome.
FT   TRANSMEM       7     27       POTENTIAL.
FT   TRANSMEM      44     64       POTENTIAL.
FT   TRANSMEM      96    116       POTENTIAL.
FT   TRANSMEM     125    145       POTENTIAL.
FT   TRANSMEM     158    178       POTENTIAL.
FT   TRANSMEM     227    247       POTENTIAL.
FT   TRANSMEM     254    274       POTENTIAL.
FT   TRANSMEM     277    297       POTENTIAL.
FT   TRANSMEM     352    372       POTENTIAL.
SQ   SEQUENCE   404 AA;  42733 MW;  93A8C960ECAD2B9C CRC64;

Query Match             18.7%;  Score 494;  DB 1;  Length 404;
Best Local Similarity   30.8%;  Pred. No. 1.7e-27;
Matches  135;  Conservative  63;  Mismatches  183;  Indels  58;  Gaps  14;

Qy   48 DNTYLLFSAYLVFSMQLGFAMLCAGSVRAKNTMNIMLTNVLDAAAGGLFYYLFGFAFAFG 107
Db    5 DTVFMFFCALLVWLWLMTPGLALFYGGMVKSKQNVLSTAMHSFSSIAIVSIVWLFGYTLAF- 63

Qy  108 SPSNGFI-GKHFFGLKDIPSSSYDYS------YFLYQWAFAIAAAGITSGSIAERTQFV 159
Db   64 APGNSIIGGLEWAGLKGVGFDPGDYSDTIPHSLFMFQMTFAVLTTAIISGAFAERMRFG 123

Qy  160 AYLIYSSFLTGFVYPVVSHWFWSPDGWASAFKITDRLFSTGVIDFAGSGVVHMVGGIAGL 219
Db  124 AFLLFSVLWASLVYTPVAHVWG-------GW--IGQLGALDFAGGNVVHISSGVAGL 173

Qy  220 WGALIEGPRMGRFDHAGRAVALRGHSAASLVVLGTFLLWFGWYGFNPGSFNKILLTYGNSG 279
Db  174 VLAIVLGKRKD------GTASSPHNLIYTFLGGALIWFGWFGFNVGS----ALTLDGVA 222

Qy  280 NYYGQWSAVGRTAVTTLLAGSTAALTTLFGKRVISGHWNVTDVCNGLLGGFAAITAGCSV 339
Db  223 MY--------AFINTNTAAAGIAGWILVEWIINKKPTMLGAVSGAIAGLVAITPAAGF 273

Qy  340 VEPWAAIVCGFVASIVLI-ACNKLABKVKFDDPLEAAQLHGGCGTWGVIFTALFAKKEYV 398
Db  274 VTPFASIIGIIGGAVCFWGVFSLKKKFGYDDALDAFGLKHGIGGIWGGIATGLFA----T 329

Qy  399 KEVYGLGRAHGLLMGGGGKL------LAAHVIQILVIAGWVSATMGPLFWGLNKLKLLRI 452
Db  330 TSVNSAG-ADGLFYGDASLIWKQIVAIAATYVFVFIVT-FVIKIVSLFLP-----LRA 381

Qy  453 SSEDELAGMDMTRHGGFAY 471
```

```
Db  382 TEEESLGLDLTMHGEKAY 400

RESULT 15
YD43_METJA
ID   YD43_METJA     STANDARD;     PRT;    420 AA.
AC   Q58739;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Putative ammonium transporter MJ1343.
GN   MJ1343.
OS   Methanococcus jannaschii.
OC   Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC   Methanocaldococcaceae; Methanocaldococcus.
OX   NCBI_TaxID=2190;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX   MEDLINE=96337999; PubMed=8688087;
RA   Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA   Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,
RA   Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA   Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA   Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA   Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA   Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA   Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT   "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT   jannaschii.";
RL   Science 273:1058-1073(1996).
CC   -|- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC   -|- SIMILARITY: Belongs to the ammonium transporter (TC 2.A.49)
CC       family.
CC   ----------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
DR   EMBL; U67574; AAB99352.1; -.
DR   PIR; F64467; F64467.
DR   TIGR; MJ1343; -.
DR   InterPro; IPR001905; Ammonium_transpt.
DR   Pfam; PF00909; Ammonium_transp; 1.
DR   TIGRFAMs; TIGR00836; amt; 1.
DR   PROSITE; PS01219; AMMONIUM_TRANSP; 1.
KW   Hypothetical protein; Transport; Transmembrane; Complete proteome.
FT   TRANSMEM      34     54       POTENTIAL.
FT   TRANSMEM      71     91       POTENTIAL.
FT   TRANSMEM     120    140       POTENTIAL.
FT   TRANSMEM     149    169       POTENTIAL.
FT   TRANSMEM     180    200       POTENTIAL.
FT   TRANSMEM     220    240       POTENTIAL.
FT   TRANSMEM     250    270       POTENTIAL.
FT   TRANSMEM     273    293       POTENTIAL.
FT   TRANSMEM     295    315       POTENTIAL.
FT   TRANSMEM     339    359       POTENTIAL.
FT   TRANSMEM     365    385       POTENTIAL.
SQ   SEQUENCE   420 AA;  43785 MW;  F7AE2554D19DACA2 CRC64;

Query Match             18.4%;  Score 485;  DB 1;  Length 420;
Best Local Similarity   30.5%;  Pred. No. 7.7e-27;
Matches  145;  Conservative  60;  Mismatches  193;  Indels  78;  Gaps  14;

Qy   16 GPNTTDASAAASLICGHFAAVDSKFVDTAFAVDNTYLLFSAYLLVFSMQLGFAMLCAGSVR 75
Db    3 GENIATADLFANATDIH--SIVQALTTLANASDVFFLVVMGVLVFMMQWGFAMLEGGQVR 60

Qy   76 AKNTMNIMLTNVLDAAAGGL-FYYLFGFAFAFGSPSNGFIG--KHFFGLKDIPSSSYDYS 132
```

```
Db       61 KKNVNNVMKNMVDWLIGCVAWLPIGGILCSKGFDLSAFIDWWKQILG-TNWPNNGLDLA 119
            ||  |::|::   |::  ::|        ||                  : |::|    |
Qy      133 YPLYQWAFAIAAAGITSGSIAERTQFVAYLIYSSFLTGFVYPV---VSHWFWSPDGWASA 189
            :     |:  :|||:   ||: :  ||:   :  :||:  |
Db      120 SWFFGLVFCATAATIVSGGVAERIKFSAYVLISLIITGLLYPLFVYLGPWGASIVPWH-- 177

Qy      190 FKITDRLFSTGVIDFAGSGVVHMVGGIAGLWGALIEGPRMGRFDHAGRAVALRGHSASLV 249
            :|||| |||  :||   |  | :|||:|
Db      178 ----------DYAGSLVVHGLGGFLALGAIAALGPRIGRFVD-GRPVPILGHNIPMA 223

Qy      250 VLGTFLLWFGWYGFNPGSFNKILLTYGNSGNYYGQWSAVGRTAVTTLAGSTAALTTLFG 309
            | | ||||||||||      :                ||:|: : |
Db      224 VFGAFALAIGWYGFNVGS---SLALGDIS--------GLVCATTMAMAGGGIGALIA 270

Qy      310 KRVISGHWNVTDVCNGLLGGFRAITAGCSVVEPWAAIVCGFVASIVLIACNKLAEKVKFD 369
            |:||     :|  ::  :|  |:|       |: :  |||||||||
Db      271 SRN-----DVLFTANGIVAGLVAICSGTDVVSPIGGLIIGLIAGLQVPIVYKLVEKAGLD 325

Qy      370 DPLEAAQLHGGCGTWGVIFTALFAKKEYVKEVYGLGRAHGLLMGGGGKLLAAHVIQIL-- 427
            :||  :|: :  ||: :  : | :      :|
Db      326 DVCGVVPVHGTAGVIGAILTGILGLK-----IFG--------GAGGVSLIDQIIGAVFC 371

Qy      428 --------VIAGWVSATMGPLFWGLNKLKLLRISSBDELAGMDMTRHGGFAYAYE 474
                    :|:|        |  | |: |:
Db      372 IIYGTGLGYILAKIVGIALGG--------LRVSEEEKMGLDMAEHKMPAYPEE 417
```

Search completed: March 9, 2004, 11:54:37
Job time : 20 secs

OM protein - protein search, using sw model

          Copyright   GenCore version 5.1.6
                       Copyright (c) 1993 - 2004 Compugen Ltd.

Run on: March 9, 2004, 11:52:48 ; Search time 46 Seconds
                                         (without alignments)
                                         3429.546 Million cell updates/sec

Title: US-10-033-109-4
Perfect score: 2639
Sequence: 1 MSLPACPAEQLAQLLGPNTT.........HGMQLRRVGPNASSTPTTDE 500

Scoring table: BLOSUM62
               Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   SPTREMBL_25:*
               1: sp_archea:*
               2: sp_bacteria:*
               3: sp_fungi:*
               4: sp_human:*
               5: sp_invertebrate:*
               6: sp_mammal:*
               7: sp_mhc:*
               8: sp_organelle:*
               9: sp_phage:*
              10: sp_plant:*
              11: sp_rodent:*
              12: sp_virus:*
              13: sp_vertebrate:*
              14: sp_unclassified:*
              15: sp_rvirus:*
              16: sp_bacteriap:*
              17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB	ID	Description
1	2254.5	85.4	501	10	Q9FVI4	Q9fvi4 lotus japon
2	2249.5	85.2	502	10	Q9FSH3	Q9fsh3 lotus japon
3	2001	75.6	518	10	Q7Y1B9	Q7y1b9 lotus japon
4	1995	75.6	519	10	Q9FUH7	Q9fuh7 brassica na
5	1988.5	75.4	496	10	Q9LK16	Q9lk16 arabidopsis
6	1962	74.3	498	10	Q7XQ12	Q7xq12 oryza sativ
7	1960	74.3	532	10	Q947N1	Q947n1 oryza sativ
8	1840.5	69.7	495	10	Q947M9	Q947m9 oryza sativ
9	1792.5	67.9	533	10	O04400	O04400 oryza sativ
10	1655.5	62.7	497	10	Q947N0	Q947n0 oryza sativ
11	1463	55.4	335	10	Q92PM8	Q92pm8 nepenthes a
12	1089.5	41.3	579	10	Q8LRM5	Q8lrm5 chlamydomon
13	1087.5	41.2	579	10	Q8LRM4	Q8lrm4 chlamydomon
14	1078	40.8	539	10	Q8RUT6	Q8rut6 chlamydomon
15	968	36.7	542	10	Q8LJU0	Q8lju0 chlamydomon
16	921	34.9	511	10	Q94KI7	Q94ki7 cylindrothe
17	830	31.5	431	5	Q9BLG3	Q9blg3 dictyosteli
18	798.5	30.3	185	10	Q9SP25	Q9sp25 brassica na
19	771.5	29.2	435	16	Q9WYN3	Q9wyn3 thermotoga
20	754	28.6	431	5	Q8MXY0	Q8mxy0 dictyosteli
21	743	28.2	427	5	Q7ZIM2	Q7zim2 dictyosteli
22	717	27.2	558	5	Q17663	Q17663 caenorhabdi
23	715.5	27.1	498	16	Q8YY62	Q8yy62 anabaena sp
24	711.5	27.0	468	17	O28528	O28528 archaeoglob
25	705	26.7	498	16	Q7UGU7	Q7ugu7 rhodopirell
26	702	26.6	485	16	Q8AAC1	Q8aac1 bacteroides
27	694.5	26.3	441	16	Q8KDQ9	Q8kdq9 chlorobium
28	683	25.9	481	16	Q8DHH4	Q8dhh4 synechococc
29	680.5	25.8	456	16	Q8ERT9	Q8ert9 oceanobacil
30	677	25.7	484	2	Q9R6T8	Q9r6t8 synechococc
31	676.5	25.6	492	16	Q7V4T8	Q7v4t8 prochloroco
32	673	25.5	486	16	Q7V330	Q7v330 prochloroco
33	673	25.5	493	2	Q9REX3	Q9rex3 prochloroco
34	665	25.2	518	16	Q8YY61	Q8yy61 anabaena sp
35	659	25.0	469	17	Q8TIE5	Q8ties methanosarc
36	657	24.9	469	17	Q8PYA6	Q8pya6 methanosarc
37	655.5	24.8	651	16	Q9VFA9	Q9vfa9 drosophila
38	655	24.8	421	16	Q9K695	Q9k695 bacillus ha
39	650.5	24.6	502	16	Q7VDS3	Q7vds3 prochloroco
40	650.5	24.6	507	16	Q8DMD2	Q8dmd2 synechococc
41	647	24.5	781	16	Q8F074	Q8f074 leptospira
42	644.5	24.4	591	5	Q7ZIM1	Q7zim1 anopheles g
43	636.5	24.1	490	16	Q7U9K3	Q7u9k3 synechococc
44	617.5	23.4	491	2	Q8GAA3	Q8gaa3 synechococc
45	611.5	23.2	491	2	Q93IP6	Q93ip6 synechococc

ALIGNMENTS

```
RESULT 1
Q9FVI4
ID   Q9FVI4        PRELIMINARY;       PRT;       501 AA.
AC   Q9FVI4;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Putative ammonium transporter AMT1.1.
OS   Lotus japonicus.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids I; Fabales; Fabaceae; Papilionoideae; Loteae; Lotus.
OX   NCBI_TaxID=34305;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Nitrogen-fixing nodules;
RA   Simon U., Udvardi M.K.;
RT   "cDNA encoding a putative ammonium transporter from Lotus japonicus.";
RL   Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF182188; AAG24944.1; -.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0008519; F:ammonium transporter activity; IEA.
DR   GO; GO:0006810; P:transport; IEA.
DR   InterPro; IPR001905; Ammonium_transpt.
DR   InterPro; IPR001064; Crystallin.
DR   Pfam; PF00909; Ammonium_transp; 1.
DR   TIGRFAMs; TIGR00836; amt; 1.
DR   PROSITE; PS01219; AMMONIUM_TRANSP; 1.
DR   PROSITE; PS00225; CRYSTALLIN_BETAGAMMA; 1.
SQ   SEQUENCE 501 AA; 53324 MW; EE65AB790SD36B70 CR664;

Query Match             85.4%;    Score 2254.5;    DB 10;    Length 501;
Best Local Similarity   84.0%;    Pred. No. 3.8e-152;
Matches 421; Conservative 38; Mismatches 39; Indels 3; Gaps 2;

Qy    1 MSLPACPAEQLAQLLGPNTTDASAAASLICGHFAAVDSKFPVDTAFAVDNTYLLFSAYLVF 60
Db    1 MALPECSAANLAQLIGPNATDAAAVAGFICDQFTAVGQRFSDTAFAVDSTYLLFSAYLVF 60
```

```
Qy   61  SMQLGFAMLCAGSVRAKNTNMILTNVLDAAAGGLFYYLFGFAFAGSPSNGFIGKHFFG  120
             |||||||||||||||||||||||||||||||||||||||||||||   |||| ||||
Db   61  SMQLGFAMLCAGSVRAKNTNMILTNVLDAAAGGLFYYLFGFAFAGSNGFIGRHFFG  120

Qy  121  LKDIPSSSYDYSYFLYQWAFAIAAAGITSGSIAERTQFVAYLIYSSFLTGFVYPVVSHWF  180
             ||| : : :|||||||||||||||||||||||||||||||||||||||||||||||
Db  121  LKDVPTVAFDYSYFLYQWAFAIAAAGITSGSIAERTQFVAYLIYSSFLTGFVYPVVSHWF  180

Qy  181  WSPDGWASAFKITDRLFSTGVIDFAGSGVVHMVGGIAGLWGALIEGPRMGRFDHAGRAVA  240
             || |||||| :  ||| |||||||||||||||||| ||||||||||| ||||||||| 
Db  181  WSGDGWASATNTGNLLFGTGVIDFAGSGVVHMVGGVAGLWGALIEGPRIGRFDHAGRAVS  240

Qy  241  LRGHSASLVVLGTFLLWFGWYGFNPGSFNKILLTYGNSGNYYGQWSAVGRTAVTTTLAGS  300
             |||||||||||||| ||||||||||||| |:  :|:|||||||||| |||||||||||
Db  241  LRGHSASLVVLGTFMLWFGWYGFNPGSFTKILSAY-DSGNYYGQWSAIGRTAVTTTLAGC  299

Qy  301  TAALTTLFGKRVISGHWNVTDVCNGLLGGFAAITAGCSVVEPWAAIVCGFVASIVLIACN  360
             |||||||||||:||||||||||||||||||||||||||||||||||||||::|||||||
Db  300  TAALTTLFGKRIISGHWNVTDVCNGLLGGFAAITAGCSVVEPWAAIVCGFVAAVVLIACN  359

Qy  361  KLAEKVKFDDPLEAAQLHGGCGTWGVIFTALFAKKEYVKEVY--GLGRAHGLLMGGGGKL  418
             :|||||: |||||||||||||||| |||||||| ||| :|  |     |:||| |||||
Db  360  MLAEKVRYDDPLEAAQLHGGCGAWGIIFTALFAKEEYVNQVYPGKPGRPYGLFMGGGGKL  419

Qy  419  LAAHVIQILVIAGWSATMGPLFWGLNKLKLLRISSEDELAGMDMTRHGGFAYAYEDDET   478
             | |||||||||:||||||||||| :||:|||||:|||||||| ||||||||||||||:
Db  420  LGAHVIQILVIIGWVSATMGPLFFILNKMKLLRISTEDELAGMDLTRHGGFAYAYEDDES   479

Qy  479  HKHGMQLRRVGPNASSTPTTD  499
             || :||| :  ||:|||: |:
Db  480  HKPGIQLRKIEPNSSSTPSAE  500
```

RESULT 2
Q9FSH3
ID   Q9FSH3   PRELIMINARY;   PRT;   502 AA.
AC   Q9FSH3;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Ammonium transporter (AMT1.1).
GN   AMT1.1.
OS   Lotus japonicus.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids I; Fabales; Fabaceae; Papilionoideae; Loteae; Lotus.
OX   NCBI_TaxID=34305;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21297198; PubMed=11404021;
RA   Salvemini F., Marini AM., Riccio A., Patriarca E.J., Chiurazzi M.;
RT   "Isolation of a member of the Lotus japonicus AMT1 family.";
RL   Gene 270:237-243(2001).
DR   EMBL; AJ279059; CAC10555.1; -.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0008519; F:ammonium transporter activity; IEA.
DR   GO; GO:0006810; P:transport; IEA.
DR   InterPro; IPR001905; Ammonium_transpt.
DR   InterPro; IPR001064; Crystallin.
DR   Pfam; PF00909; Ammonium_transp; 1.
DR   TIGRFAMs; TIGR00836; amt; 1.
DR   PROSITE; PS01219; AMMONIUM_TRANSP; 1.
DR   PROSITE; PS00225; CRYSTALLIN_BETAGAMMA; 1.
SQ   SEQUENCE   502 AA; 53395 MW; BE5AB79066B159B CRC64;

Query Match   85.2%; Score 2249.5; DB 10; Length 502;
Best Local Similarity   84.0%;
Matches 420; Conservative   38; Mismatches   39; Indels   3; Gaps   2;

```
Qy    2  SLPACPAEQLAQLLGPNTTDASAAASLICGHFAAVDSKFVDTAFAVDNTYLLFSAYLVFS   61
             : | ||:|||||::| | ||:|:|   |||||:||   |:||||||:||||||||||||
Db    3  ALPECSAANLAQLIGNATDAAVAGPICDQFTAVGQRFSDTAFAVDSTYLLFSAYLVFS    62
```

```
Qy   62  MQLGFAMLCAGSVRAKNTNMILTNVLDAAAGGLFYYLFGFAFAGSPSNGFIGKHFFGL  121
             |||||||||||||||||||||||||||||||||||||||||||||   |||| ||||
Db   63  MQLGFAMLCAGSVRAKNTNMILTNVLDAAAGGLFYYLFGFAFAGSNGFIGRHFFGL  122

Qy  122  KDIPSSSYDYSYFLYQWAFAIAAAGITSGSIAERTQFVAYLIYSSFLTGFVYPVVSHWFW  181
             :||: : :|||||||||||||||||||||||||||||||||||||||||||||||||
Db  123  KDVPTVAFDYSYFLYQWAFAIAAAGITSGSIAERTQFVAYLIYSSFLTGFVYPVVSHWFW  182

Qy  182  SPDGWASAFKITDRLFSTGVIDFAGSGVVHMVGGIAGLWGALIEGPRMGRFDHAGRAVAL  241
             | |||||| :  ||| |||||||||||||||||| ||||||||||| ||||||||| |
Db  183  SGDGWASATNTGNLLFGTGVIDFAGSGVVHMVGGVAGLWGALIEGPRIGRFDHAGRAVSL  242

Qy  242  RGHSASLVVLGTFLLWFGWYGFNPGSFNKILLTYGNSGNYYGQWSAVGRTAVTTTLAGST  301
             |||||||||||||| ||||||||||||| |:  :|:|||||||||| |||||||||||
Db  243  RGHSASLVVLGTFMLWFGWYGFNPGSFTKILSAY-DSGNYYGQWSAIGRTAVTTTLAGCT  301

Qy  302  AALTTLFGKRVISGHWNVTDVCNGLLGGFAAITAGCSVVEPWAAIVCGFVASIVLIACNK  361
             |||||||||||:||||||||||||||||||||||||||||||||||||||::|||||||
Db  302  AALTTLFGKRIISGHWNVTDVCNGLLGGFAAITAGCSVVEPWAAIVCGFVAAVVLIACNM  361

Qy  362  LAEKVKFDDPLEAAQLHGGCGTWGVIFTALFAKKEYVKEVY--GLGRAHGLLMGGGGKLL  419
             |||||: |||||||||||||||| |||||||| ||| :|  |     |:||| ||||||
Db  362  LAEKVRYDDPLEAAQLHGGCGAWGIIFTALFAKEEYVNQVYPGKPGRPYGLFMGGGGKLL  421

Qy  420  AAHVIQILVIAGWSATMGPLFWGLNKLKLLRISSSEDELAGMDMTRHGGFAYAYEDDETH  479
             ||||||||||:||||||||||| :||:|||||::|||||||| ||||||||||||||:
Db  422  GAHVIQILVIIGWVSATMGPLFFINKMKLLRISTEDELAGMDLTRHGGFAYAYEDDESH  481

Qy  480  KHGMQLRRVGPNASSTPTTD  499
             || :||| :  ||:|||: |:
Db  482  KPGIQLRKIEPNSSSTPSAE  501
```

RESULT 3
Q7Y1B9
ID   Q7Y1B9   PRELIMINARY;   PRT;   518 AA.
AC   Q7Y1B9;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Ammonium transporter.
GN   AMT1;2.
OS   Lotus japonicus.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids I; Fabales; Fabaceae; Papilionoideae; Loteae; Lotus.
OX   NCBI_TaxID=34305;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   El Alaoui H., Chiurazzi M.;
RT   "Characterization of a new ammonium transporter in Lotus japonicus.";
RL   Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY135020; AAM95453.1; -.
SQ   SEQUENCE   518 AA; 54830 MW; BE5526F05105SC0 CRC64;

Query Match   75.8%; Score 2001; DB 10; Length 518;
Best Local Similarity   75.8%; Pred. No. 4.2e-134;
Matches 383; Conservative   41; Mismatches   73; Indels   8; Gaps   5;

```
Qy    1  MSLPACPAEQLAQLLGPNT-TDASAAASLICGHFAAVDSKFVDTAFAVDNTYLLFSAYLV   59
             |: :|| | ||||||||  :|| ||: |||:|| |||||| ||| ||||||||||||||
Db    1  MASLSCSATDLAPLLTATTNATATAAATYLCNQLDTISRKLSDTTYAVDNTYLLFSAYLV   60

Qy   60  FSMQLGFAMLCAGSVRAKNTMNIMLTNVLDAAAGGLFYYLFGFAFAGSPSNGFIGKHFF  119
             | |||||||||||||||||||||||||||||||||||| ||||||||||| |
Db   61  FAMQLGFAMLCAGSVRAKNTMNIMLTNVLDAAAGGLSYYLFGFAFAFAGAPSNGFIGRHFF  120

Qy  120  GLKDIPSSSYDYSYFLYQWAFAIAAAGITSGSIAERTQFVAYLIYSSFLTGFVYPVSHW  179
             ||  : :  :||||  ||||||||||||||||||||||||||||||||||||||||||
Db  121  GLKHYPSPTYDYSFFLYQWAFAIAAAGITSGSIAERTQFVAYLIYSSFLTGFVYPVVSHW  180

Qy  180  FWSPDGWASAFKIT-DRLFSTGVIDPAGSGVVHMVGGIAGLWGALIEGPRMGRFDHAGRA  238
             |   | |  :  :   : :|| |:   |::
Db  ...
```

```
Db   181  LWSSDGWASPTRTGSLLFGSGAIDFAGSGVVHVGGIAGLWGAPIEGPRIGRFPDRTGRS   240
Qy   239  VALRGHSASLVVLGTFLLWFGWYGFNPGSFNKILLTYGN---SGNYYGQWSAVGRTAVTT  295
Db   241  VALRGHSASLVVLGSFLLMFGWYGFNPGSFLTIAKAYGNNGENGNYYGQWSAIGRTAVTT  300
Qy   296  TLAGSTAALTTLFGKRVISGHWNVTDVCNGLLGGFAAITAGCSVVEPWAAIVCGFVASIV  355
Db   301  TLAGCTAALTTTLFSKRLLEGHWKVLDVCNGLLGGFAAITSGCSVVEPWPAIVCGFVAAWV 360
Qy   356  LIACNKLAEKVKFDDPLEAAQLHGGCGTWGVIFTALFAKKEYKEVYG-LGRAHGLLMGG   414
Db   361  LIGLNLVAAKMKYDDPLEAAQLHGGCGAWGVLFTGLFAKGEYVAEVYGSAGRPYGLLMGG  420
Qy   415  GGKLLAAHVIQIIVIAGWVSATMGPLFWGLNKLKLLRISSEDELAGMDWTRHGGFAYAYE  474
Db   421  GGKLLAAQVIEILVVCGWTATMGPLFYGLHKTKJLRISEDDETAGMDLTRHGGFAYAYH   480
Qy   475  DDE--THKHGMQLRRVGPNASSTPT                                    497
Db   481  DDDVSTKRGVMMSRIGPGSSSPST                                     505
```

RESULT 4
Q9FUH7
ID   Q9FUH7;   PRELIMINARY;   PRT;   519 AA.
AC   Q9FUH7;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   High-affinity ammonium transporter AMT1,2.
OS   Brassica napus (Rape).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Brassicales; Brassicaceae; Brassica.
OX   NCBI_TaxID=3708;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Pearson J.N., Finnemann J., Schjoerring J.K.;
RT   "Regulation of the high-affinity ammonium transporter (BnAMT1;2) in
RT   the leaves of Brassica napus by nitrogen status.";
RL   Submitted (SEP-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF306518; AAG28780.1; -.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0008519; F:ammonium transporter activity; IEA.
DR   GO; GO:0006810; P:transport; IEA.
DR   InterPro; IPR001905; Ammonium transpt.
DR   Pfam; PF00909; Ammonium_transp; 1.
DR   TIGRFAMs; TIGR00836; amt; 1.
DR   PROSITE; PS01219; AMMONIUM_TRANSP; 1.
SQ   SEQUENCE   519 AA; 55759 MW; 2F30322B1A146D4E CRC64;

Query Match   75.6%;   Score 1995;   DB 10; Length 519;
Best Local Similarity 77.1%;   Pred. No. 1.1e-133;
Matches 378;   Conservative 38; Mismatches 62; Indels 12; Gaps 5;

```
Qy     6  CPAEQLAQLLGPNTTDASAAASLICGHFAAVDSKFVDTAFAVDNTYLLFSAYLVFSMQLG   65
Db     7  CSAADLATLLGPN--ATAAADYICGQLGTVNNKFTDAAFAIDNTYLLFSAYLVFAMQLG   63
Qy    66  FAMLCAGSVRAKNTMNIMLTNVLDAAAAGGLFYLFGFAPFGSPSNGFIGKHFFGLKDIP  125
Db    64  FAMLCAGSVRAKNTMNIMLTNVLDAAAGGFFYYLFGYAPAFGSSEGFIGRHFALRDFP   123
Qy   126  SSSYDYSFLYQWAFAIAAAGITSGSIAERTQFVAYLIYSSFLTGFVYPVVSHWFWSPDG  185
Db   124  TPTADYSFFLYQWAFAIAAAGITSGSIAERTQFVAYLIYSSFLTGFVYPVVSHWFWSPDG 183
Qy   186  WASAFK-ITDRLFSTGVIDFAGSGVVHMVGGIAGLWGALIEGPRMGRFDHAGRAVALRGH 244
Db   184  WASPFRSADDRLFSTGAIDFAGSGVVHMVGGIAGLWGALIEGPRRGRFEKGGRAIALRGH 243
Qy   245  SASLVVLGTFLLWFGWYGFNPGSFNKILLTYGNSGNYYGQWSAVGRTAVTTLAGSTAAL  304
```

```
Db   244  SASLVVLGTFLLWFGWYGFNPGSFTKILVPY-NSGSNYGQWSGIGRTAVNTTLSGCTAAL 302
Qy   305  TTLFGKRVISGHWNVTDVCNGLLGGFAAITAGCSVVEPWAAIVCGFVASIVLIACNKLAE 364
Db   303  TTLFGKRLLSGHWNVTDVCNGLLGGFAAITAGCSVVEPWAAIVCGFMTSLVLLGCNKLAE 362
Qy   365  KVKFDDPLEAAQLHGGCGTWGVIFTALFAKKEYKEVYGL--GRAHGLLMGGGKLLAAH   422
Db   363  LVQYGDPLEAAQLHGGCGAWGLIFVGLFAKEKYLNEVYGATPGRPYGLFMGGGGKLLGAQ 422
Qy   423  VIQILVIAGWVSATMGPLFWGLNKLKLLRISSEDELAGMDWTRHGGFAYAYEDDETHKHG 482
Db   423  LVQIIVIVGWVSATMGTLFFILKRLNLLRISEQHEMQGMDMTRHGGFAYIYHDNNDDESH- 481
Qy   483  MQLRRVGPNA                                                   492
Db   482  ----RVDPGS                                                   487
```

RESULT 5
Q9LK16
ID   Q9LK16;   PRELIMINARY;   PRT;   496 AA.
AC   Q9LK16;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Ammonium transporter.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Columbia;
RA   Kaneko T., Kato T., Sato S., Nakamura Y., Asamizu E., Tabata S.;
RL   Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Columbia;
RX   MEDLINE=20363099; PubMed=10907853;
RA   Nakamura Y.;
RT   "Structural analysis of Arabidopsis thaliana chromosome 3. II.
RT   Sequence features of the regions of 4,251,695 bp covered by ninety P1,
RT   TAC and BAC clones.";
RL   DNA Res. 7:217-221(2000).
DR   EMBL; AP000382; BAB02928.1; -.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0008519; F:ammonium transporter activity; IEA.
DR   GO; GO:0006810; P:transport; IEA.
DR   InterPro; IPR001905; Ammonium transpt.
DR   Pfam; PF00909; Ammonium transp; 1.
DR   TIGRFAMs; TIGR00836; amt; 1.
DR   PROSITE; PS01219; AMMONIUM_TRANSP; 1.
SQ   SEQUENCE   496 AA; 52638 MW; FB97C38A94D17B5B CRC64;

Query Match   75.4%;   Score 1988.5; DB 10; Length 496;
Best Local Similarity 76.1%;   Pred. No. 3.1e-133;
Matches 375;   Conservative 46; Mismatches 65; Indels 7; Gaps 4;

```
Qy     6  CPAEQLAQLLGPNTTDASAAASLICGHFAAVDSKFVDTAFAVDNTYLLFSAYLVFSMQLG  65
Db     7  CSAADLSALLGPN--ATAAADYICGQLGSVNNKFTDAAYAIDNTYLLFSAYLVFAMQLG   63
Qy    66  FAMLCAGSVRAKNTMNIMLTNVLDAAAAGGLFYLFGFAPFGSPSNGFPIGKHFGLKDIP  125
Db    64  FAMLCAGSVRAKNTMNIMLTNVLDAAAGGLFYLFGYAPAFGESSDGFIGRHNFGLQNFP  123
Qy   126  SSSYDYSFLYQWAFAIAAAGITSGSIAERTQFVAYLIYSSFLTGFVYPVVSHWFWSPDG  185
Db   124  TLTSDYSFFLYQWAFAIAAAGITSGSIAERTKFVAYLIYSSFLTGFVYPVVSHWFWSPDG 183
```

```
Qy   242  RGHSASLVVLGTFLLWFGWYGFNPGSFNKILLTYGNSGNYYGQWSAVGRTAVTTTLAGST  301
          :||||||||||||||||||| |||||| ::|||||||||| ||:  :|  :  ||||||||
Db   230  KGHSASLVVLGTFLLWFGWYGFNPGSFTTILKTYGPAGGINGQWSGVGRTAVTTTLAGSV  289

Qy   302  AALTTLFGKRVISGHWNVTDVCNGLLGGFAAITAGCSVVEPWAAIVCGFVASIVLIACNK  361
          |||||||||||||||| ||: |||| ||||||||||||||  ::||||  :|||||| :|
Db   290  AALTTLFGKRLQTGHWNVVDVCNGLLGGFAAITAGCSVVDPWAAIICGFVSAWVLIGLNA  349

Qy   362  LAEKVKFDDPLEAAQLHGGCGTWGVIFTALFAKKEYVKEVYGLGRAHGLLMGGGKLLAA   421
          ||: :||||||||||||||||||||||||| ||:::|:::|  |||||||||||||||||
Db   350  LAARLKFDDPLEAAQLHGGCGAWGLLFTALFARQKYVEEIYGAGRPYGLFMGGGKLLAA   409

Qy   422  HVIQILVIAGWVSATMGPLFWGLNKLKLRISSEDELAGMDTRHGGFAYAYEDDETHKH   481
          ||||||| ||||:||||||||||||||:||||:||:|||||||||||||::: |:|||:
Db   410  HVIQILVIFGWVSCTMGPLFYGLXKLGLLRISAEDETSGMDLTRHGGFAYVVHDEDEHDX  469

Qy   482  ----GMQLR----RVGPNASSTPTTD  499
                :|::    || ||||||||:
Db   470  SGVGGFMLRSAQTRVEPAAAAASNSN  495
```

```
RESULT 7
ID   Q947N1
AC   Q947N1;
DT   01-DEC-2001  (TrEMBLrel. 19, Created)
DT   01-DEC-2001  (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2003  (TrEMBLrel. 24, Last annotation update)
DE   Ammonium transporter 1-1.
GN   OSAMT1-1.
OS   Oryza sativa (Rice).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=4530;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Hoque M.S., Masle J., Udvardi M.K., Upadhyaya N.M.;
RT   "Cloning and characterization of three Ammonium transporter genes from
RT   rice.";
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF289477; AAL05612.1; -.
DR   Gramene; Q947N1; -.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0008519; F:ammonium transporter activity; IEA.
DR   GO; GO:0006810; P:transport; IEA.
DR   InterPro; IPR001905; Ammonium_transpt.
DR   Pfam; PF00909; Ammonium_transp; 1.
DR   TIGRFAMs; TIGR00836; amt; 1.
DR   PROSITE; PS01219; AMMONIUM_TRANSP; 1.
SQ   SEQUENCE  532 AA;  56774 MW;  E3BF7656B9A6C890 CRC64;

Query Match             74.3%;  Score 1960;  Pred. No. 3.5e-131;  Length 532;
Best Local Similarity   75.2%;
Matches  376;  Conservative  43;  Mismatches  61;  Indels  20;  Gaps   6;

Qy     3  LPACPAEQLAQLLGPNTTDASAAASLICGHFAAVDSKFVDTAFAVDNTYLLFSAYLVFSM   62
          ::|:: :||||||||||||||||||||:|||||||||||      :||||||||||||||
Db     1  MATCAAD-LAPLLGP--VAANATDYLCNRFA------DTTSAVDATYLLFSAYLVFAM    49

Qy    63  QLGFAMLCAGSVRAKNTMNIMLTNVLDAAAGGLFYYLFGFAFAFGSPSNGFIGKHFFGLK  122
          |||||||||||||||||||||||||||||||:|||||||||:||||||||||||:||||||
Db    50  QLGFAMLCAGSVRAKNTMNIMLTNVLDAAAGALFYYLFGFARAFGTPSNGFIGKQFFGLK  109

Qy   123  DIPSSSYDYSFYLYQWAFAIAAAGITSGGSIAERTQFVAYLIYSSFLTGFVYPVVSHWFWS  182
          ::|: :||:|||:||||| ||||||||||||||||||||:|||:|||:||||| ||| |
Db   110  HMPQTGFDYDFFLFQWAFAIAAAGITSGGSIAERTQFVAYLIYSAFLTGFVYPVVSHWIWS  169

Qy   183  PDGWASAFKIT-DRLFSTGVIDFAGSGVVHMVGGIAGLWGALIEGPRMGRFDHAGRAVAL  241
          ||||||:|| |||||||||||||||||||||| |||||||||||||||||:|||||:|||
Db   170  ADGWASASRTSGPLLFGSGVIDFAGSGVVFHMVGGVAGLWGALIEGPRIGRFDHAGRSVAL  229
```

Qy   242 RGHSASLVVLGTFLLWFGWYGFNPGSFNKILLTYGNSGNYYGQWSAVGRTAVTTTLAGST 301
               :||||||||||||||||||||||||||||||| ||||||:||||||||||||||||||
Db   230 KGHSASLVVLGTFLLWFGWYGFNPGSFTTILKTYGPAGINGQWSGVGRTAVTTTLAGSV 289

Qy   302 AALTTLFGKRVISGHWNVTDVCNGLLGGFAAITAGCSVVEPWAAIVCGFVASIVLIACNK 361
               ||||||||||:||||| :|||||||||||||||||||||||||||||::||||: |
Db   290 AALTTLFGKRLQTGHWNVVDVCNGLLGGFAAITAGCSVVDPWAAIICGFVSAWVLIGLNA 349

Qy   362 LAEKVKFDDPLEAAQLHGGCGTWGVIFTALFAKKEYVKEVYGLGRAHGLLMGGGGKLLAA 421
               :|||||||||||||||||||||||||||||||::||||| :||||||||||||||
Db   350 LAARLKFDDPLEAAQLHGGCGAWGILFTALFARQKYVEEIYGAGRPYGLFMGGGGKLLAA 409

Qy   422 HVIQILVIAGWVSATMGPLFWGLNKLKLLRISSEDELAGMDMTRHGGFAYAYEDDETHKH 481
               ||||||:||||||||||||||||||||||||||||||||||||||||| |:||
Db   410 HVIQILVIFGWVSCTMGPLFYGLKKLGLLRISAEDETSGMDLTRHGGFAYVYHDEDEHDK 469

Qy   482 ----GMQLR----RVGPNAS 493
               ||| :|
Db   470 SGVGGFMLRSAQTRVEPAAA 489

RESULT 8

Q947M9

ID   Q947M9   PRELIMINARY;   PRT;   495 AA.
AC   Q947M9;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Ammonium transporter 1-3.
GN   OSAMT1-3.
OS   Oryza sativa (Rice).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=4530;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Hoque M.S., Masle J., Udvardi M.K., Upadhyaya N.M.;
RT   "Cloning and characterization of three Ammonium transporter genes from
RT   rice.";
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF289479; AAL05614.1; -.
DR   Gramene; Q947M9; -.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0008519; F:ammonium transporter activity; IEA.
DR   GO; GO:0006810; P:transport; IEA.
DR   InterPro; IPR001905; Ammonium_transpt.
DR   Pfam; PF00909; Ammonium_transp; 1.
DR   TIGRFAMs; TIGR00836; amt; 1.
SQ   SEQUENCE   495 AA; 52524 MW; 5BB3914F6741AE10 CRC64;

Query Match   69.7%; Score 1840.5; DB 10; Length 495;
Best Local Similarity   75.6%; Pred. No. 1e-122;
  Matches 357; Conservative 38; Mismatches 64; Indels 13; Gaps 5;

Qy   8 AEQLAQLLGPNTDASAAASLICGHFAAVDSKFVDTAFAVDNTYLLFSAYLVFSMQLGPA 67
               |:| |||||||| |: || |||| || |||||||||||||||||||||||||||||
Db   5 ADTLGPLLG---TAAAANTDYLCNQFA------DTTSAVDSTYLLFSAYLVFAMQLGFA 54

Qy   68 MLCAGSVRAKNTMNIMLTNVLDAAAGGLFYLFGPAFGSPSNGFIGKHFFGLKDIPSS 127
               ||||||||||||||||||||||||||||||||||||||:|||||||||||:| :|
Db   55 MLCAGSVRAKNTMNIMLTNVLDAAAGALFYLFGFAFAFGSPSNGFIGKHFFGLKQVPQV 114

Qy   128 SYDYSYFLYQWAFAIAAAGITSGSIAERTQFVAYLIYSSFLTGFYPVVSHWFWSPDGWA 187
               ||||||||||||||||||||||||:|||||||||||||||||| ::|||: ||||||
Db   115 GFDYSFFLFQWAFAIAAAGITSGSIAERTQFVAYLIYSAFLTGFVPVVSHWIWSADGWA 174

Qy   188 SAFKIT-DRLFSTGVIDFAGSGVHMVGGIAGLWGALIEGPRMGRFDHAGRAVALRGHSA 246
               ||| | ||||||||||||||||||||||||||||||||||:|||:|||||||||||||
Db   175 SASRTSGSLLFGSGVIDFAGSGVHMVVAACRTL-GRPHRGSPHWRFDHAGRSVALRGHSA 233

Qy   247 SLVVLGTFLLWFGWYGFNPGSFNKILLTYGNSGNYYGQWSAVGRTAVTTTLAGSTAALTT 306
               ||||||||||||||||||||||| ||| || :: ||
Db   247 ...

Db   234 SLVVLGSFLLWFGWYGFNPGSFLILKSYGPPGSIHGQWSAVGRTAVTTTLAGSTAALTT 293

Qy   307 LFGKRVISGHWNVTDVCNGLLGGFAAITAGCSVVEPWAAIVCGFVASIVLIACNKLAEKV 366
               ||||:| :||| :||||||||||||||||| :|||:|:|:|| ::| ||:
Db   294 LFGKRLQTGHWNVDVCNGLLGGFAAITAGCSVVDPWAAIICGFVSAWVLIGLNALAAARL 353

Qy   367 KFDDPLEAAQLHGGCGTWGVIFTALFAKKEYVKEVYGL-GRAHGLLMGGGGKLLAAHVIQ 425
               |||||||||||||||| :|||||||||:|:||:::: | |:|||||||:|||||||||||
Db   354 KFDDPLEAAQLHGGCGAWGVIFTALFARKEYVDQIFGQPGRPYGLFMGGGGKLLGAHIVV 413

Qy   426 ILVIAGWVSATMGPLFWGLNKLKLLRISSEDELAGMDMTRHGGFAYAYAEDDE 477
               ||||||||||||||||||||:|||:||||::|:|:|:||:||||||||:||||
Db   414 ILVIAAWVSFTMAPLFLVLNKLGLLRISAEDEMAGMDQTRHGGFAYAYHDDD 465

RESULT 9

O04400

ID   O04400   PRELIMINARY;   PRT;   533 AA.
AC   O04400;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Putative ammonium transporter OsAMT1p.
GN   OSAMT1.
OS   Oryza sativa (Rice).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=4530;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Nipponbare; TISSUE=Root;
RX   MEDLINE=94341258; PubMed=8062823;
RA   Ninnemann O., Jauniaux J.C., Frommer W.B.;
RT   "Identification of a high affinity NH4+ transporter from plants.";
RL   EMBO J. 13:3464-3471(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Nipponbare; TISSUE=Root;
RA   von Wiren N., Bergfeld A., Ninnemann O., Frommer W.B.;
RT   "OsAMT1-1, a putative ammonium transporter cDNA from rice.";
RL   Submitted (APR-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF001505; AAB58937.1; -.
DR   PIR; T03441; T03441.
DR   Gramene; O04400; -.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0008519; F:ammonium transporter activity; IEA.
DR   GO; GO:0006810; P:transport; IEA.
DR   InterPro; IPR001905; Ammonium_transpt.
DR   Pfam; PF00909; Ammonium_transp; 1.
DR   TIGRFAMs; TIGR00836; amt; 1.
DR   PROSITE; PS01219; AMMONIUM_TRANSP; 1.
SQ   SEQUENCE   533 AA; 56943 MW; AA06F72D188A0E44 CRC64;

Query Match   67.9%; Score 1792.5; DB 10; Length 533;
Best Local Similarity   71.6%; Pred. No. 2.9e-119;
  Matches 361; Conservative 39; Mismatches 77; Indels 27; Gaps 11;

Qy   3 LPACPAEQLAQLLGPNTDASAAASLICGHFAAVDSKFVDTAFAVDNTYLLFSAYLVFSM 62
               | || |||| || |||| |: || |||| |||||||||||||||||||||||||||
Db   1 MATCAAD-LAPLLGP---VAANATDYLCNRFA-------DTTSAVDATYLLFSAYLVFAM 49

Qy   63 QLGPAMLCAGSVRAKNTMNIMLTNVLDAAAGGLFYLFGPAF--AFGSPSNGFIGKHFF 119
               |||| |||||||||||||||||||||||||||:|||||| | |||:|||||||||:||
Db   50 QLGFAMLCAGSVRAKNTMNIMLTNVLDAAAGALFYLFGFASLRDCLRTPSNGFIGKQFF 109

Qy   120 GLKDIPS-SSYDYSYFLYQWAFAIAAAGITSGSIAERTQFVAYLIYSSFLTGFYPVVSH 178
               |||:||: |||||||||||||||||||||||||||||||||||||||:|||||||||||
Db   110 GLKMPARPGSDYDFFLFQWAFAIAAAGITSGSIAERTQFVAYLIYSAFLTGFVPVVSH 169

Qy   179 WFWSPDGWASAFKIT-DRLFSTGVIDFAGSGVHMVGGIAGLWGALIEGPRMGRFDHAGR 237
               | || ||||| || | ||||||||||||||||||||||||||||||||||:|||:||||
Db   170 WIWSADGWASASRTSGPLLFGSGVIDFAGSGVHMVVGGVAGLWGALIEGPRIGRFDHAGR 229

```
Qy  246 ASLVVLGTFLLWFGWYGFNPGSFNKILLTYGNSGNYYGQWSAVGRTAVTTTLLAGSTAALT 305
Db  235 ASLVVLGTFLLWFGWYGFNPGSFTTISKIYGESGTIDGQWSAVGRTAVTTSLAGSVAAL-  293

Qy  306 TLFGKRVISGHWNVTDVCNGLL--------GGFAAITAGCSVVEPWAAIVCGFVASIV    355
Db  294 --------NHAVRQEMADGALERDRRLQRSPRRVRAITAGCSVVDPWASVICGFVSAWV   344

Qy  356 LIACNKLAEKVKFDDPLEAAQLHGGCGTWGVIFTALFAKKEYVKEVYGL-GRAHGLLMGG  414
Db  345 LIGCNKLALMLKFDDPLEATQLHGGCGAWGIIFTALFARKEYVELIYGVPGRPYGLFMGG  404

Qy  415 GGKLLAAHVIQILVIAGWVSATMGPLFWGLNKLKLLRISSEDELAGMDTRHGGFAYAYE   474
Db  405 GGRLLAAHIVQILVIVGWVSATMGTLFYVLHRFGLLRVSTSTEMGMDPSCHGGFGGTVDE  464

Qy  475 DDETH----KHGMQLRRVGPNAS 493
Db  465 DEGQRRVRAKSAAETARVEPRKS 487
```

RESULT 11

```
ID   Q9ZPM8          PRELIMINARY;      PRT;     335 AA.
Q9ZPM8
AC   Q9ZPM8;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Ammonium transporter (Fragment)
GN   AMT1.
OS   Nepenthes alata (Winged pitcher plant).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Caryophyllales; Nepenthaceae; Nepenthes.
OX   NCBI_TaxID=4376;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Pitcher;
RA   Schulze W., Frommer W.B., Ward J.M.;
RT   "Transporters for ammonium, amino acids and peptides are expressed in
RT   pitchers of the carnivorous plant Nepenthes.";
RL   Plant J. 17;101-110(1999).
DR   EMBL; AF080541; AAD16012.1; -.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0008519; F:ammonium transporter activity; IEA.
DR   GO; GO:0006810; P:transport; IEA.
DR   InterPro; IPR001905; Ammonium_transpt.
DR   Pfam; PF00909; Ammonium_transp; 1.
DR   TIGRFAMs; TIGR00836; amt; 1.
DR   PROSITE; PS01219; AMMONIUM_TRANSP; 1.
FT   NON_TER        1      1
FT   NON_TER      335    335
SQ   SEQUENCE   335 AA;  35653 MW;  752FD5CDFEB75E9F CRC64;

Query Match             55.4%;  Score 1463;  DB 10;  Pred. No. 4.3e-96;
Best Local Similarity   81.2%;
Matches  272;    Conservative   27;    Mismatches  34;    Indels   2;    Gaps  1;

Qy  136 YQWAFAIAAAGITSGSIAERTQFVAYLIYSSFLTGFVYPVVSHWFWSPDGWASAFKITDR 195
Db    1 YQWAFAIAAAGITSGSIAERTQFVAYLIYSSFLTGFVYPIVSHWFWSVDGWASASRASNL  60

Qy  196 LFSTGVIDFAGSGVVHMVGGIAGLWGALIEGPRMGRFDHAGRAVALRGHSASLVVLGTFL 255
Db   61 LLGGSVTDFAGSGVVHMVGGVSGLWGALIEGPRIGRFDRSGRSVILRGHSGSLVVLGTFL 120

Qy  256 LWFGWYGFNPGSFNKILLTYGNSGNYYGQWSAVGRTAVTTTLLAGSTAALTTLFGKRVISG 315
Db  121 LWFGWYGFNPGSFLTILKAYGNNGGYFGQWSAVGRTAVTTTLLAGSTAALTTLFGKRLLVG 180

Qy  316 HWNVTDVCNGLLGGFAAITAGCSVVEPWAAIVCGFVASIVLIACNKLAEKVKFDDPLEAA 375
Db  181 HWNVTDVCNGLLGGFAAITSGCSVVDPWAAIICGFVSSWVLIGFNKLAEKVKYDDPLEAA 240
```

```
Qy  238 AVALRGHSASLVVLGTFLLWFGWYGFNPGSFNKILLTYGNSGNYYGQWSAVGRTAVTTTL 297
Db  230 SVALKGHSASLVVLGTFLLWFGWYGFNPGSFTTILKTYGPAGGINGQWSGVGRTAVTTTL 289

Qy  298 AGSTRALLTTLFGKRVISGHWNVTDVCNGLLGGFAAITAGCSVVEPWAAIVCGFVASIVLi 357
Db  290 AGSVAALTTLFGKRLQTGHWNVVDVCNGLLGGFAAITAGCGVVDPWAAIICGFVSAWVLi  349

Qy  358 ACNKLAEKVKFDDPLEAAQLHGGCGTWGVIFTALFAKKEYVKEVYGLGRAHGLLMGGGGK 417
Db  350 GLN-LAARLKFDDPLEAAQLHGGCGAWGILFTALFARQKYV-ERSTAPAAYGLFM-GRRK  406

Qy  418 LLAAHVIQILVIAGWVSATMGPLFWGLNKLKLLRISSEDELAGMDTRHGGFAYAYEDDE  477
Db  407 LLVGAVIQILVIFGWVSCTMGPLFYGLKKLGLLRISGRDETSGMDLTRHGGFAYVYHDED 466

Qy  478 THKH----RVGPNAS 493
Db  467 EHDKSGVGGFMLRSAQTRVEPAAA 490
```

RESULT 10

```
ID   Q947N0          PRELIMINARY;      PRT;     497 AA.
Q947N0
AC   Q947N0;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Ammonium transporter 1-2.
GN   OSAMT1-2.
OS   Oryza sativa (Rice).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=4530;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Hoque M.S., Masle J., Udvardi M.K., Upadhyaya N.M.;
RT   "Cloning and characterization of three Ammonium transporter genes from
RT   rice.";
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF289478; AAL05613.1; -.
DR   Gramene; Q947N0; -.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0008519; F:ammonium transporter activity; IEA.
DR   GO; GO:0006810; P:transport; IEA.
DR   InterPro; IPR001905; Ammonium_transpt.
DR   Pfam; PF00909; Ammonium_transp; 1.
DR   TIGRFAMs; TIGR00836; amt; 1.
DR   PROSITE; PS01219; AMMONIUM_TRANSP; 1.
SQ   SEQUENCE   497 AA;  53398 MW;  8A054402C44C3D24 CRC64;

Query Match             62.7%;  Score 1655.5;  DB 10;  Pred. No. 1.4e-109;
Best Local Similarity   65.6%;
Matches  330;    Conservative   53;    Mismatches  81;    Indels  39;    Gaps  9;

Qy    9 EQLAQLIG--PNTTDASAAASLICGHFAAVDSKFVDTAFAVDNTYLLFSAYLVFSMQLGF  66
Db    6 DSLGPLIGGGAANSTD--AANYIC------NRFTDTSSAVDATYLLFSAYLVFAMQLGF   55

Qy   67 AMLCAGSVRAKNTMNIMLTNVLDAAAGGLFYILFGPAFAFGSPSNGFPIGKCHFFGLKDIPS 126
Db   56 AMLCAGSVRAKNSMNIMLTNVFDAAAGALFYYLFGPFA-SRRTPSKGFPIGKQFFGLKHMPQ 114

Qy  127 SSYDYSYFLYQWAFAIAAAGITSGSIAERTQFVAYLIYSSFLTGFVYPVVSHWFWSPDGW 186
Db  115 TGYDYDFFLPQWAFAIAAAGITSGSIAERTRFSAYLIYSAFLTGFVYPVVSHWFWSTDGW 174

Qy  187 ASAFKIT-DRLFSTGVIDFAGSGVVHMVGGIAGLWGALIEGPRMGRFDHAGRAVALRGHS 245
Db  175 ASAGRLLTGPLLFKSGVIDFAGSGVVHLVGGIAGLWGAFIIEGPRIGRFDAAGRTVAMKGHS 234
```

```
Qy  376  QLHGGCGTWGVIFTALFAKKEYKEVYG--LGRAHGLLMGGGGKLLAAHVIQILVIAGWV  433
         ||||||| ||:|||| |||| ||||  :  ||| |||||||||| ||:||:||:|| |||
Db  241  QLHGGCGMWGLIFTGLFATKKKYVSEVYGGSPDRPHGLFMGGGGKLLAAQLIQIVVIFGWV 300

Qy  434  SATMGPLFWGLNKLKLLRISSEDELAGMDMTRHGG  468
         :|||||||||||| || ||| :|   |||||||||
Db  301  TATMGPLFYGLKWLKLLRVKSEHEMAGMDMTSHGG  335
```

RESULT 12
Q8LRM5
ID   Q8LRM5   PRELIMINARY;   PRT;   579 AA.
AC   Q8LRM5;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Putative amt protein.
OS   Chlamydomonas reinhardtii.
OC   Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC   Chlamydomonadaceae; Chlamydomonas.
OX   NCBI_TaxID=3055;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CC-125;
RA   Yaoi T., Feild E., King N., Soupene E., Kim K.-S., Kustu S.;
RT   "Chlamydomonas reinhardtii amt gene.";
RL   Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF509496; AAM43910.1; -.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0008519; F:ammonium transporter activity; IEA.
DR   GO; GO:0006810; P:transport; IEA.
DR   InterPro; IPR001905; Ammonium_transpt.
DR   Pfam; PF00909; Ammonium_transp; 1.
DR   TIGRFAMs; TIGR00836; amt; 1.
DR   PROSITE; PS01219; AMMONIUM_TRANSP; 1.
SQ   SEQUENCE   579 AA; 60683 MW; B20C4CCA09FB5C2B CRC64;

```
Query Match             41.3%;
Best Local Similarity   44.7%;                  Score 1089.5;  DB 10;  Length 579;
Matches  235;  Conservative  74;  Mismatches 144;  Indels  73;  Gaps  16;   Pred. No. 3e-69;

Qy    5  ACPAEQL---AQLLGPNTDASAAASLIGH----FAAVDS--KFV--------------   41
         :  : ||   : ||:|                 : :||:   ::
Db   16  SCTDEQMLEVTELVGD-----SATAAIVCSQYDCSFTAVSCVLKYLAAQQADLSEQMDGT  70

Qy   42  ---DTAFAVDNTYLLFSAYLVFS-MQLGFAMLCAGSVRAKNTMNIMLTNVLDAAAGGLFY   97
            : ::|| ||:||||||:|  ||||||::|||:||:|| ||:||: ||||:    | |
Db   71  AKAEVALSLDVAFLLFSAYLVFGPMQLGFALLCAGAIRSKNSMNVLMKNILDACTGAIGF  130

Qy   98  YLFGFAPAFG-----SPSNGFIGKHFFGL------KDIPSS-SYD------YSYFLYQWAF  140
         ||||:| |||     ||||:|||::       :  |||:                |||:|
Db  131  YLFGYAFAFGHHANQTSNAFIGDHNFALSYTTQVSSLDSNVSYDGFATQGWHVFFFQWSF  190

Qy  141  AIAAAGITSGSIAERTQFVAYLIYSSFLTGFVYPVVSHWFWSPDGWASAFKITDR-----  195
          ::|  ||:|:::||| |:|||:|::|: |||||||: ||:|||:||:|:|
Db  191  CARAATTIVSGAVAERCTFQAYLAYAFFISSFVYPVVVHWVWSASGWLSAFN-TSRDGYAL 249

Qy  196  LFSTGVIDFAGSGVVHMVGGIAGLWGALIEGPRMGRFDHAGRAVALRGHSASLVVLGTFL  255
         |: ||:||||| ||||:  |:| |||: | ||:|||| | : |:  |||::|||:|||||
Db  250  LLQTGAIDFAGGGVVHLTGMAALMGAWIIGPRIGRFDASGKVNEMKGHSATLVVMGTFL  309

Qy  256  LWFGWYGFNPGSFNKILLTYGNSGNYYGQWSAVGRTAVTTLAGSTAALTTLFGKRVISG  315
         ||||:|||||| |  :  ::::   : :     |:|||||| ||:|| |:|| :: :
Db  310  LWFGFYGFNPGS-NLTIATTASA------IVSRVAVTTTLSAGAAGLTGLFWRYMRTS  361

Qy  316  HWNVTDVCNGLLGGFAAITAGCSVVEPWAAIVCGFVASIVLIACNKLA-EKVKFDDPLEA  374
          |:||:|||  |||: | ||  ||||||||::|||||::|:   | |  |||||| |:.|
Db  362  TWDTVLVCNCCLAGLVGITCSCSVVEPWAALICGFVAAFVFIGFEYVVLYKMKIDDPVSA  421

Qy  375  AQLHGGCGTWGVIFTALFAKKEYKEVYGL--------GRAHGLLMGGGGKLLAAHVIQ  425
         . ||  ||.||::| .| | : :  | |         |||:|||:|||| ||| :|:|
Db  422  VALHLFCGVWGLLFPGLLAQPTYVADVGAYGFGPDVKGSKKFGILYGGHGQVLLCQVIE  481
```

```
Qy  426  ILVIAGWVSATMGPLFWGLNKLKLLRISSEDELAGMDMTRHGGFAY  471
         .|.| .|:|:||| |:||||:|.|:|:| ||::|:|:|:||: :|:
Db  482  ALSICAWVGVMGAFFGLLKVAKRLRVPVDQELAGLDVSKHTNEAY   527
```

RESULT 13
Q8LRM4
ID   Q8LRM4   PRELIMINARY;   PRT;   579 AA.
AC   Q8LRM4;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Putative amt protein.
OS   Chlamydomonas reinhardtii.
OC   Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC   Chlamydomonadaceae; Chlamydomonas.
OX   NCBI_TaxID=3055;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C-9;
RA   Yaoi T., Feild E., King N., Soupene E., Kim K.-S., Kustu S.;
RT   "Chlamydomonas reinhardtii amt gene.";
RL   Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF509497; AAM43911.1; -.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0008519; F:ammonium transporter activity; IEA.
DR   GO; GO:0006810; P:transport; IEA.
DR   InterPro; IPR001905; Ammonium_transpt.
DR   Pfam; PF00909; Ammonium_transp; 1.
DR   TIGRFAMs; TIGR00836; amt; 1.
DR   PROSITE; PS01219; AMMONIUM_TRANSP; 1.
SQ   SEQUENCE   579 AA; 60697 MW; 170C49CF09FB5C2D CRC64;

```
Query Match             41.2%;
Best Local Similarity   44.7%;                  Score 1087.5;  DB 10;  Length 579;
Matches  235;  Conservative  74;  Mismatches 144;  Indels  73;  Gaps  16;   Pred. No. 4.1e-69;

Qy    5  ACPAEQL---AQLLGPNTDASAAASLIGH----FAAVDS--KFV--------------   41
         :  : ||   : ||:|                 : :||:   ::
Db   16  SCTDEQMLEVTELVGD-----SATAAIVCSQYDCSFTAVSCVLKYLAAQQADLSEQMDGT  70

Qy   42  ---DTAFAVDNTYLLFSAYLVFS-MQLGFAMLCAGSVRAKNTMNIMLTNVLDAAAGGLFY   97
            : ::|| ||:||||||:|  ||||||::|||:||:|| ||:||: ||||:    | |
Db   71  AKAEVALSLDVAFLLFSAYLVFGPMQLGFALLCAGAIRSKNSMNVLMKNILDACTGAIGF  130

Qy   98  YLFGFAPAFG-----SPSNGFIGKHFFGL------KDIPSS-SYD------YSYFLYQWAF  140
         ||||:| |||     ||||:|||::       :  |||:                |||:|
Db  131  YLFGYAFAFGHHANQTSNAFIGDHNFALSYTTQVSSLDSNVSYDGFATQGWHVFFFQWSF  190

Qy  141  AIAAAGITSGSIAERTQFVAYLIYSSFLTGFVYPVVSHWFWSPDGWASAFKITDR-----  195
          ::|  ||:|:::||| |:|||:|::|: |||||||: ||:|||:||:|:|
Db  191  CARAATTIVSGAVAERCTFQAYLAYAFFISSFVYPVVVHWVWSASGWLSAFN-TSRDGYAL 249

Qy  196  LFSTGVIDFAGSGVVHMVGGIAGLWGALIEGPRMGRFDHAGRAVALRGHSASLVVLGTFL  255
         |: ||:||||| ||||:  |:| |||: | ||:|||| | : |:  |||::|||:|||||
Db  250  LLQTGAIDFAGGGVVHLTGMAALMGAWIIGPRIGRFDASGKVNEMKGHSATLVVMGTFL  309

Qy  256  LWFGWYGFNPGSFNKILLTYGNSGNYYGQWSAVGRTAVTTLAGSTAALTTLFGKRVISG  315
         ||||:|||||| |  :  ::::   : :     |:|||||| ||:|| |:|| :: :
Db  310  LWFGFYGFNPGS-NLTIATTASA------IVSRVAVTTTLSAGAAGLTGLFWRYMRTS  361

Qy  316  HWNVTDVCNGLLGGFAAITAGCSVVEPWAAIVCGFVASIVLIACNKLA-EKVKFDDPLEA  374
          |:||:|||  |||: | ||  ||||||||::|||||::|:   | |  |||||| |:.|
Db  362  TWDTVLVCNCCLAGLVGITCSCSVVEPWAALICGFVAAFVFIGFEYVVLYKMKIDDPVSA  421

Qy  375  AQLHGGCGTWGVIFTALFAKKEYKEVYGL--------GRAHGLLMGGGGKLLAAHVIQ  425
         . ||  ||.||::| .| | : :  | |         |||:|||:|||| ||| :|:|
Db  422  VALHLFCGVWGLLFPGLLAQPTYVADVGAYGFGPDVKGSKKFGILYGGHGQVLLCQVIE  481

Qy  426  ILVIAGWVSATMGPLFWGLNKLKLLRISSEDELAGMDMTRHGGFAY  471
         .|.| .|:|:||| |:||||:|.|:|:| ||::|:|:|:||: :|:
Db  482  ALSICAWVGVMGAFFGLLKVAKRLRVPVDQELAGLDVSKHTNEAY   527
```

```
Qy    439  PLFWGLNKLKLRISSEDELAGMDMTRH 466
           |||| ||| ::|||| :::||:|
Db    482  SFFFILNKAGLLRVPLQBEMAGLDAANY 509
```

RESULT 15
Q8LJU0
```
ID   Q8LJU0;
AC   Q8LJU0;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Putative ammonium transport 1,2.
OS   Chlamydomonas reinhardtii
OC   Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC   Chlamydomonadaceae; Chlamydomonas.
OX   NCBI_TaxID=3055;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Camargo A., Gonzalez-Ballester D., Fernandez E.;
RT   "cDNA for a putative ammonium transport from Chlamydomonas reinhardtii
RT   (CrAmt 1;2).";
RL   Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF530051; AAM94623.1; -.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0008519; F:ammonium transporter activity; IEA.
DR   GO; GO:0006810; P:transport; IEA.
DR   InterPro; IPR001905; Ammonium_transpt.
DR   Pfam; PF00909; Ammonium_transp; 1.
DR   TIGRFAMs; TIGR00836; amt; 1.
SQ   SEQUENCE  542 AA;  57596 MW;  830942DED136E2D8 CRC64;
```

```
Query Match             36.7%;
Best Local Similarity   42.6%;    Score 968;  DB 10;  Pred. No. 1.2e-60;   Length 542;
Matches  222;  Conservative  66;  Mismatches  165;  Indels  68;  Gaps  14;
```

```
Qy      5  ACPAEQLAQLLGPNTTDASAAASLICGHFAA-----------------VD----SKFVD------ 42
           |||  ::|:|                         ||              ||
Db      4  ACLLAAKSALVG--VANGDSLASAICAAMEAAAPVARHLQETTVDLTTLQBTTVDTIGLNL 61

Qy     43  --TAFAVDNTYLLFSAYLVFSMQLGFAMLCAGSVRAKNTMNIMLTNVLDAAAGGLFYLF 100
             :|: ||||:|||  ||||||||||||||||| |||||||||||| ||: ||||
Db     62  QGTMDGLNTLLLYGGALVFLMHGGSAMLEAGAIRSKNAVNILLQTVLDASCSALMYFV 121

Qy    101  GFAFAF--GSPSNGFIGKHFFGLKDIP-----SSSYDYSFLYQWAFAIAAAGITSGSIA 153
           |||||:  |::|||  |:::|||  :          :::|| |||  |||: ::|:
Db    122  GFAFAFGIGDKPNKFIGNAMFALVDIDTHNTGSGTGKWIDWFFQWAFAATAVTIPAGAVA 181

Qy    154  ERTQFVAYLIYSSFLTGFVYPVVSHWFWSPDGWASAFKITDRLFSTGVIDFAGSGVVHMV 213
           |::||  ||||:|||:||||||| |||||  :|:|||||: ||| || ||||||||
Db    182  ERRFNFAYLGYSMFIGGWVYPIIAHWVWCIEGWL-GYGVIKPFINAGMIDFAGSGVIHMT 240

Qy    214  GGIAGLWGALIEGPRMGRFDHAGRAVALRGHSASLVVLGTFLLWFGWYGFNPGSFNKILL 273
           |||||| || ||||:||||  ||  ||||||  ||:|||:||||||||||||    :||
Db    241  GGLAGLIGAIMVDPRLGRFDADGKPVDMPGHSALLVVLGTVLLWFGWYGFNPAS---VLL 297

Qy    274  TYGNSGNYYGQWSAVGRTAVTTLLAGSTAALTTI-FGKRVISGH--WNVTDVCNGLLGGF 330
           ||||:|||| || ||||||||  ||||:|      :|| |:     |||||| |||| |
Db    298  I--NSSTYA--IVCGRAAVCTTLLAGAAGGVSCLIFG---FARHRGWDLVGLCNGILCGF 349

Qy    331  AAITAGCSVVEPWAAIVCGFVASIVLIACNKLAEKVKFDDPLEAAQLHGGCGTWGVIFTA 390
           |||:||||| ||||||||||||| ::|||  ||  |||:||||  |:|||| |:|||:||
Db    350  VAVTACPHVIYPWAAIICGLVAGLWFEFLCWLLLKLKIDDPLSAGPMHFGCCMWGVFFTG 409

Qy    391  LFAKKEYVKEVYGLG-----------------RAHGLLMGGGKLLAAHVIQILVIA 430
           |:|::||||||| |                  ||||||||||  ||||||||||:|:
Db    410  LFARQQYIQEYYVHGPSAAGELSKFYPWGAFYKGYNADGELVMSDGKLLASQIVGILVII 469

Qy    431  GWYSATMGPLFWGLNKLKLLRISSEDELAGMDMTRHGGFAY 471
           ||: | |||||||||| :||| ||||||||||||  |:||
Db    470  GWVIGMVVPFFGIFQFFGALRIPPEEEMGLDRSKHGGSAY 510
```

RESULT 14
Q8RUT6
```
ID   Q8RUT6;
AC   Q8RUT6;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Putative ammonium transporter.
GN   AMT1-1.
OS   Chlamydomonas reinhardtii.
OC   Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC   Chlamydomonadaceae; Chlamydomonas.
OX   NCBI_TaxID=3055;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Gonzalez-Ballester D., Perez-Alegre M., Fernandez E.;
RT   "A putative ammonium transporter (CrAMT1-1) from Chlamydomonas
RT   reinhardtii.";
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Gonzalez-Ballester D., Fernandez E.;
RT   "A putative ammonium transporter cDNA of Chlamydomonas reinhardtii.";
RL   Submitted (JAN-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY058211; AAL38652.1; -.
DR   EMBL; AF479643; AAL85345.1; -.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0008519; F:ammonium transporter activity; IEA.
DR   GO; GO:0006810; P:transport; IEA.
DR   InterPro; IPR001905; Ammonium_transpt.
DR   Pfam; PF00909; Ammonium_transp; 1.
DR   TIGRFAMs; TIGR00836; amt; 1.
DR   PROSITE; PS01219; AMMONIUM_TRANSP; 1.
SQ   SEQUENCE  539 AA;  57147 MW;  8281E2BBA33E1DD2 CRC64;
```

```
Query Match             40.8%;
Best Local Similarity   43.9%;    Score 1078;  DB 10;  Pred. No. 1.8e-68;   Length 539;
Matches  223;  Conservative  68;  Mismatches  165;  Indels  52;  Gaps  8;
```

```
Qy      3  LPACPAEQLAQLLGPNTTDASAAA-----------------SLICGHFAAVDSKFVDTA 44
           | ||| ::|:|                           :  ||| || ||:||:|
Db     10  LGSCSVETVTALLGYGLEQDSITALCQPEGGAGCTSTDNCMFQYLMGATADAQSTASDVG 69

Qy     45  FAVDNTYLLFSAYLVFSMQLGFAMLCAGSVRAKNTMNIMLTNVLDAAAGGLFYLFGFAF 104
           ||||:||||||||||||||||||||||||||| ||||||||||||||||||||| ||||
Db     70  VGLDVSFLLFSGYLVFVMQLGFAVLCAGSIRSKNCMNILLKNMLDACVGAIGFVFGYAF 129

Qy    105  A----FGSPSNGFIGKHFFGLKDIPSSS-------YDYSFLYQWAFAIAAAGITSGSI 152
           :     |::|||  |:::|||    :            :::||  ||||||| |||
Db    130  AYGRKYGQNSNGFIGNWNFALSYTTQISMSGTEFTTFGWHQFFQWSFCAATTIVSGAV 189

Qy    153  AERTQFVAYLIYSSFLTGFVYPVVSHWFWSPDGWASAFKITD----RLFSTGVIDFAGSG 208
           ||||| |||||:||||:||||||| |||||  ||:|||| |     ||:||| |||||||
Db    190  AERCTFMAYMIYAFFLSSFVYPVIVHWTWDGQGWLSAFNTFQDGYALILKTGAIDFAGSG 249

Qy    209  VVHMVGGIAGLWGALIEGPRMGRFDHAGRAVALRGHSASLVVLGTFLLWFGWYGFNPGSF 268
           |||| ||:|||| || |||:|||    || |||||| ||:|||||:|||||||||||||
Db    250  VVHMTGIAALMGAWIMGPRVGRFANDGTVNEMRGHSSTLVVMGTFLLWFGWYGFNPGS- 308

Qy    269  NKILLTYGNSGNYYGQWSAVGRTAVTTTLAGSTAALTTLFGKRVISGHWNVTDVCNGLLG 328
           || |||||:|||| || ||  |||  |||||| :| ||   | :||  || || |||:|
Db    309  NLVVASQAAA------TVVSRVAVTTALIAGGAGGISMLFYKFLTVKAWDVVATCNGILA 361

Qy    329  GFAAITAGCSVVEPWAAIVCGFVASIVL-IACNKLAEKVKFDDPLEAAQLHGGCGTWGVI 387
           |  |:||||| ||||||| ::||  |:   |:|  ||  |||:|||    ||||| |:|
Db    362  GLVAVTASCSVIEPWAAIITGAIGAIIFSIADYVTLYKLKVDDPVSAFALHGAVGAWGVL 421

Qy    388  FTALFAKKEYVKEVYGL--------GRAHGLLMGGGKLLAAHVIQILVIAGWYSATMG 438
           :  |:|::||||||| |         |||||||||| |||||||||||:|:||: | |
Db    422  FPGFLAAPHYVVEVYGAYGFGMDAREGKRFGLFYGGHGGVLLVQLIEVLAIFGWTGFMMG 481
```